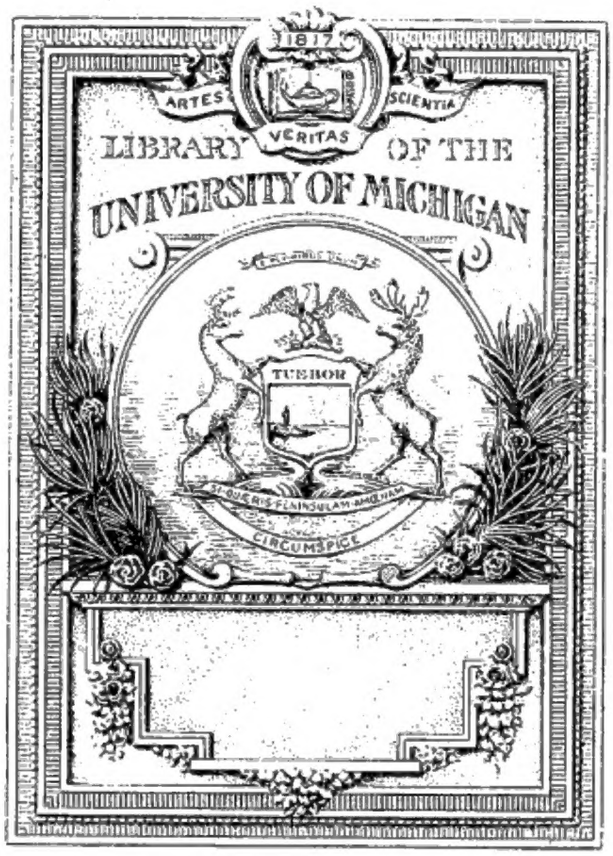
B 3 9015 00229 187 3
University of Michigan – BUHR


1817
ARTES
SCIENTIA
VERITAS
LIBRARY OF THE
UNIVERSITY OF MICHIGAN
TUEBOR
CIRCUMSPICE

BF
431
.584
1920

Die Intelligenz der Kinder und Jugendlichen

und die Methoden ihrer Untersuchung

An Stelle einer dritten Auflage des Buches:
Die Intelligenzprüfung an Kindern
und Jugendlichen

Von

William Stern

Verlag von Johann Ambrosius Barth / Leipzig 1920

Copyright by Johann Ambrosius Barth in Leipzig, 1920.
Alle Rechte, insbesondere das der Übersetzung, vorbehalten.

Druck von Grimme & Trömel in Leipzig.

Meinen Mitarbeitern!

391009

transfer
to Gen. Lib
7. 11. 64

Vorwort.

Das vorliegende Buch ist aus kleinen Anfängen hervorgegangen. Im Jahre 1912 hatte ich für den V. Psychologenkongreß in Berlin ein Sammelreferat über die Methodik der Intelligenzprüfungen übernommen. Dieses Referat wurde, um es pädagogischen und ärztlichen Kreisen zugänglich zu machen, als Sonderausgabe veröffentlicht unter dem Titel „Die psychologischen Methoden der Intelligenzprüfung und deren Anwendung an Schulkindern" (106 S.). Als im Jahre 1916 eine Neuausgabe nötig wurde, beschränkte ich mich auf anastatischen Neudruck des alten Textes und fügte einen zweiten Teil an, der die inzwischen eingetretenen Fortschritte zum Gegenstand hatte. Diese „Fortschritte auf dem Gebiete der Intelligenzprüfung 1912—1915" (64 S.) wurden für Besitzer der ersten Ausgabe auch gesondert herausgegeben.

Ein solches Verfahren ließ sich aber nicht ein zweites Mal wiederholen; das erneute Bedürfnis einer nochmaligen Ausgabe verlangte vielmehr eine tiefgreifende Neugestaltung. Denn in den letzten Jahren waren gerade auf dem Gebiet der Intelligenz-Forschung und -Prüfung gänzlich neue Probleme und Gesichtspunkte hervorgetreten, die unsere psychologische Arbeit teils gründlich veränderten, teils bedeutend erweiterten. Theoretisch bedurften die Fragen nach Wesen, Arten und Entwicklung der Intelligenz, die im früheren Buch nur andeutend behandelt wären, einer genaueren Erörterung. Methodisch war inzwischen die einseitige Vorherrschaft des Binet-Simon-Verfahrens überwunden worden; insbesondere hatten die neuen Probleme der Schülerauslese und der Berufseignung die Ausbildung andersartiger Methoden gefordert. Und zwar erstreckte sich diese Methodenentwicklung nicht nur auf die Neuausbildung experimenteller Verfahren (vor allem für die höheren Intelligenzfunktionen und die praktische Intelligenz),

sondern sie führte auch zu der Einsicht, daß das Experiment durch beobachtende Methoden ergänzt werden müsse, und so wurden Beobachtungsbogen und Intelligenzschätzung zu neuen Hilfsmitteln unserer Forschung. Ferner mußte eine empfindliche Lücke der älteren Arbeit ausgefüllt werden: es war eine Übersicht über die einzelnen Intelligenzfunktionen und die Art ihrer Untersuchung zu geben. Endlich hatte sich inzwischen der Bestand an Ergebnissen beträchtlich vermehrt, die teils den internationalen Massenuntersuchungen mit der Binet-Methode, teils anderen Verfahrungsweisen zu danken waren; sie beziehen sich auf die Verteilung der Intelligenzen, auf ihre Beschaffenheit bei Schwachbefähigten und Hochbefähigten, auf ihre Abhängigkeit von sozialen Bedingungen, auf ihr Verhältnis zur eigentlichen Schulleistungsfähigkeit.

Um diesen Aufgaben genügen zu können, mußte die ursprüngliche Absicht einer Umarbeitung der alten Auflage aufgegeben und ein neues Buch geschrieben werden. Aus den früheren Auflagen sind nur wenige Textseiten und eine Reihe von Tabellen übernommen worden. Erleichtert wurde mir die Neuarbeit dadurch, daß im Hamburger Psychologischen Laboratorium während der letzten Jahre das Begabungsproblem im Mittelpunkt sowohl der theoretisch-psychologischen wie der angewandt-psychologischen Arbeit stand und noch steht. Die hier angestellten Studien (deren ausführliche Veröffentlichung in den „Hamburger Arbeiten zur Begabungsforschung“ fortlaufend erfolgt) boten mir eine breite empirische Grundlage zur Behandlung meines Themas. In besonders engem Zusammenhang mit dem vorliegenden Buch steht Nr. III der „Hamburger Arbeiten“, die eine „Methodensammlung zur Intelligenzprüfung von Kindern und Jugendlichen“, zusammengestellt von W. Stern und O. Wiegmann, enthält und für die zurzeit in Deutschland bekannten Tests Vollständigkeit anstrebt. Da dieses Buch als Ergänzung des vorliegenden zu betrachten ist, so konnte hier auf eine detaillierte Beschreibung der einzelnen Methoden verzichtet werden.

Neben den Untersuchungen des Hamburger Instituts suchte ich auch die anderweitig getane Arbeit in möglichst weitem Umfange mit heranzuziehen und zwar ebenso die psychologische Forschungsarbeit wie die zu pädagogischen Zwecken entfaltete Prüfungstätigkeit. Hierbei konnte ich mich nicht immer auf bloße Berichterstattung oder Übernahme gefundener Ergebnisse beschränken; vielmehr war an vielen Stellen Kritik notwendig. Die Neuheit

unseres Problemgebiets und seine enge Beziehung zu drängenden praktischen Lebensaufgaben bringt es mit sich, daß nicht immer der notwendige Ernst, die Objektivität und die Selbstkritik gewahrt worden ist, die man ganz besonders für angewandt-psychologische Arbeit verlangen muß. Die Kritik dieses Buches bezieht sich z. T. auf fachpsychologische Untersuchungen, Prüfungen und Hypothesen, z. T. auf die Stellungnahme der Praktiker zu unserem Problem. Lehrer, Schulbehörden und Ärzte hatten ja Anlaß, sich mit den Methoden und Ergebnissen der Intelligenzpsychologie zu beschäftigen; und da findet sich manches, was die wissenschaftliche Psychologie nicht widerspruchslos geschehen lassen kann. Bald wurde ihr das Recht abgesprochen, an der Bearbeitung praktischer, insbesondere pädagogischer Kulturaufgaben mitzuwirken; bald wurden ihre Methoden ohne genügende Vorsicht übernommen, dilettantisch mißbraucht und zu voreiligen Folgerungen verwandt, ja wohl gar tendenziös in den Dienst bestimmter schulpolitischer Überzeugungen gestellt. Indem das Buch die eine wie die andere Einseitigkeit bekämpft, sucht es zugleich die Wege zu ebnen für ein ernsthaftes und förderliches Zusammenarbeiten von Psychologie und Pädagogik.

Daß freilich das Thema unseres Buches nur ein enges Teilgebiet aus dieser Berührungssphäre von Psychologie und Pädagogik herausschneidet, muß doch vielleicht noch ausdrücklich betont werden; sonst könnte die besondere Beschäftigung mit dem Intelligenzproblem als eine Hinneigung zum Intellektualismus ausgelegt werden, und das hieße die Absicht dieser Arbeit von Grund aus mißverstehen. Man begegnet ja einer Überschätzung der Intelligenz im Rahmen der Gesamtheit des psychischen Lebens nicht selten; und gerade bei eifrigen, aber einseitigen Vertretern der Intelligenzprüfung ist zuweilen der Wahn zu finden, als käme es nur auf die Intelligenzfeststellung an, um einen Menschen richtig zu bewerten, zu beraten und zu behandeln. Diese intellektualistische Gefahr besteht aber nicht für den, der vom Standpunkt des Personalismus an die psychologischen Fragen herangeht; denn er weiß, daß im gesamten Zwecksystem der menschlichen Persönlichkeit die Intelligenz eine zwar nicht unwichtige, aber doch durchaus partiale Rolle spielt, und gerade das eingehende Studium ihres Wesens und ihrer Leistungen ist geeignet, zugleich ihre Abgrenzung gegen andere Seiten und ihre nur dienende Aufgabe gegenüber den eigentlichen Zielsetzungen persönlichen Daseins festzustellen. Die Intelligenz als solche ist immer nur Mittel und Rüstzeug (und zwar nur

eines neben anderen), niemals selber richtunggebender, wertbestimmender Zweck im Leben des Einzelmenschen und der Gemeinschaft.

Bei der technischen Fertigstellung des Buches, der Herstellung des Literaturverzeichnisses und des Sachregisters sowie bei der Korrekturlesung war mir der ständige Mitarbeiter am Hamburger Laboratorium, Herr Lehrer Otto Wiegmann, sehr hilfreich. Ihm sei hierfür herzlich gedankt.

Hamburg, am 30. Mai 1920.

W. Stern.

Folgende Abkürzungen werden verwandt:

I. = Intelligenz
IA. = Intelligenzalter
IP. = Intelligenzprüfung
IQ. = Intelligenzquotient
IR. = Intelligenzrückstand
ISch. = Intelligenzschätzung
BS. = Methode Binet-Simon
LA. = Lebensalter
SchL. = Schulleistung
Korr. = Korrelation
MS. = „Methodensammlung" von Stern und Wiegmann.

Die *kursiv* gedruckten Ziffern beziehen sich auf die Nummern des Literaturverzeichnisses.

Inhaltsverzeichnis.

Erster Teil: Wesen, Arten und Entwicklung der Intelligenz.

Seite

I. Kapitel: Begriff der Intelligenz.
1. Versuch einer Begriffsbestimmung 1
2. Andere Begriffsbestimmungen der Intelligenz 7
II. Kapitel: Die Bedingungen der Intelligenz 11
III. Kapitel: Die Verschiedenheiten der Intelligenz 18
1. Intelligenzarten und -typen 18
2. Intelligenzgrade 25
IV. Kapitel: Die Entwicklung der Intelligenz 29
1. Wachstum der Intelligenz 29
2. Wandlungen der Intelligenz 33

Zweiter Teil: Die Untersuchung der Intelligenz.

V. Kapitel: Hauptgesichtspunkte der Methodik 45
1. Allgemeines 45
2. Experiment (Test) und Beobachtung 48
3. Forschungs- und Prüfungsmethoden 53
4. Untersuchung der Einzelfunktionen und der Gesamtintelligenz 58
5. Intelligenzmaße 61
VI. Kapitel: Prüfung der einzelnen Intelligenzfunktionen . . 70
1. Kenntnisprüfungen 72
2. Prüfung der Merkfähigkeit 74
3. Prüfung des räumlichen Vorstellens 78
4. Beurteilung technisch-physikalischer Verhältnisse. (Vorstellen von Verschiebungen im Raume) 81
5. Ausführung technischer Aufgaben. Praktische Intelligenz . 82
6. Aufmerksamkeits- und Konzentrationstests 84
7. Suggestibilität 85
8. Prüfung der Sprachbeherrschung 87
9. Assoziationen 89
10. Bildbetrachtung 91
11. Erkennungstests 94
12. Verhalten zu vorgestellten Situationen 95
13. Kritikfähigkeit 100
14. Definitionen 104
15. Schließen 108

Seite

16. Ergänzungstests 110
17. Drei- und Zweiwortmethode 113
18. Finden des Wesentlichen 116
19. Vergleichen 118
20. Ordnungstests 121
21. Zuordnungstests 127

VII. Kapitel: Prüfsysteme zur Messung des Intelligenzgrades . 127
1. Psychiatrische Prüfreihen 129
2. Die Methode der Altersstaffelung nach Binet-Simon: Das Prüfungsverfahren 133
3. Die Maßprinzipien der Staffelmethode 140
4. Mängel der Staffelmethode und Änderungsvorschläge . . . 144
5. Begabungs- und Entwicklungstests. (Meumanns Vorschlag) 151
6. Die Punktskala-Methode 156

Dritter Teil: Anwendungen und Ergebnisse. Intelligenz und Schule.

VIII. Kapitel: Gesamtverteilungen der Intelligenzen 158

IX. Kapitel: Intelligenzuntersuchungen nicht normaler Kinder und Jugendlicher nach der BS.-Methode 169
1. Der diagnostische Wert der Intelligenzprüfung 169
2. Massenuntersuchungen an Hilfsschulen 178
3. Die verschiedenen Schwachsinnsgruppen 181
4. Das Verhalten zu den einzelnen Tests 186
5. Intelligenzprüfungen sittlich Minderwertiger 190

X. Kapitel: Intelligenz und Schulleistung 194
1. Klassenalter und Schulzeugnisse in ihrer Beziehung zur Intelligenz 196
2. Intelligenzschätzung 200
3. Intelligenzschätzung und Schulrangordnung 209
4. Testrangordnungen in ihren Beziehungen zu anderen Rangordnungen 217

XI. Kapitel: Die sozialen Bedingungen der Intelligenz 226
1. Vergleichende BS.-Prüfungen von Kindern verschiedener sozialer Schichten 226
2. Weitere Intelligenzvergleichungen von Kindern verschiedener sozialer Schichten 237
3. Die Ursachen der Unterschiede 242
4. Pädagogische Ausblicke 246

Vierter Teil: Psychologische Methoden der Schülerauslese.

XII. Kapitel: Methodische Grundsätze für die Mitwirkung der Psychologie an der Schülerauslese 249
1. Die Vorauslese durch den Lehrer und die Abfassung von Schülerpsychogrammen 253
2. Die Prüfung 261
3. Die Entscheidung 271
4. Erprobung und „bewegliche Front" 275

Seite

XIII. Kapitel: Übersicht über die bisherigen Schülerauslesen . . 280
1. Auslese 9jähriger Kinder (Dr. Lipmann, Berlin) 281
2. Auslese 9—11jähriger Kinder (Dr. Bobertag) 282
3. Auslese 10jähriger Kinder (Hamburg) 283
4. Aufnahmeprüfungen für die höheren Schulen („Mittelschulen") Deutsch-Österreichs (10jährige Kinder) 296
5. Prüfung 11jähriger Knaben (Göttingen) 297
6. Auslese 13jähriger Schüler und Schülerinnen (Moede und Piorkowski, Berlin) 298
7. Die Leipziger Anweisungen 304
8. Prüfung 13- und 14jähriger Volksschulkinder in Hannover . 305
9. Auslese von Handelsschülern (Beger, Leipzig) 305
10. Auslese 14- und 15jähriger Mädchen zur Aufnahme in ein Lehrerinnenseminar (Hamburg) 306
11. Ausleseprüfungen von Fachschülern und Lehrlingen der metallbearbeitenden Berufe 309

Literaturverzeichnis 313

Sachregister 329

Namenverzeichnis 332

Anhang 335

I. Teil.

Wesen, Arten und Entwicklung der Intelligenz.

I. Kapitel.

Begriff der Intelligenz.

1. Versuch einer Begriffsbestimmung.

Die neuere Psychologie, welche anfing mit dem Studium der Sinneswahrnehmungen, dann überging zu dem der Vorstellungen und Gefühle, hat nun in den letzten zwei Jahrzehnten auch die eigentlich intellektuellen Funktionen zu bearbeiten begonnen. Und hier ist es bemerkenswert, wie gleichzeitig — wenn auch ziemlich unabhängig voneinander — die allgemeine, theoretische und die differentielle, angewandte Psychologie diesen Schritt taten. Dort entwickelte sich die Psychologie des Denkens, hier die Erforschung der Intelligenzunterschiede.

Die Psychologie des Denkens beschäftigt sich mit den Denktätigkeiten und den Denkinhalten, also mit den in bestimmten Zeitpunkten vorhandenen akuten Bewußtseinsvorgängen. „Intelligenz" dagegen ist ein Ausdruck für eine Disposition, d. h. eine dauernde Fähigkeit und Strebigkeit zur Vollziehung von Denktätigkeiten; sie gehört damit gleichsam einer anderen Dimension der Persönlichkeit an[1]). Und wenn wir Intelligenzforschung treiben, Intelligenzuntersuchungen vornehmen, so ist die Absicht eben auf dieses Dispositionelle gerichtet; wir wollen die chronische geistige Beschaffenheit eines Menschen ergründen, nicht nur die in bestimmten Augenblicken vollzogenen Akte und von ihm erlebten Bewußtseinsinhalte zergliedern. So haben denn Denkpsychologie und Intelligenz-

[1]) Über den Begriff der „Disposition" und die Bedeutung, die er in der Psychologie zu spielen hat, muß auf die grundsätzlichen Ausführungen verwiesen werden, in denen ich die personalistische Theorie des psychischen Lebens entwickle. Es sind dies: Die menschliche Persönlichkeit (*16*), Die Psychologie und der Personalismus (*15*), Die menschliche Persönlichkeit und ihr psychisches Leben (*17*).

psychologie verschiedene Ziele und gehen deshalb zum Teil auch verschiedene Wege; aber selbstverständlich besteht eine gegenseitige Abhängigkeit. Die geistigen Akte und Inhalte, mit denen sich die Denkpsychologie beschäftigt, sind für den Intelligenzforscher wichtig als die Betätigungsformen der geistigen Veranlagung, die er studieren will, als Symptome, an denen er Grad und Art der vorhandenen Intelligenz erkennt und mißt; umgekehrt muß der Denkpsychologe, wenn er nicht ganz im Phänomenologischen stecken bleiben will, stets die Disposition zum Denken nach Grad und Art als Erklärungsgrund für die von ihm festgestellten Denkvorgänge selbst heranziehen.

Der oft zu hörende Einwand, daß das Problem der Intelligenzuntersuchung überhaupt nicht erfolgreich bearbeitet werden könne, ehe nicht die Phänomenologie der Denkvorgänge erschöpfend behandelt sei, ist meines Erachtens unzutreffend. In der Wissenschaft gibt es nicht ein solch sauberes Nacheinander verschiedener Forschungsaufgaben, sondern nur ein gegenseitiges Sichstützen, Beeinflussen und Ineinandergreifen. Diese Gegenseitigkeit kann und muß freilich noch in weit höherem Maße verwirklicht werden, als es bisher der Fall war; und so wird auch unsere Darstellung so manche denkpsychologische Betrachtung nachholen, die früheren Behandlungen des Themas fehlte.

Aber freilich, die eigentliche Begriffsbestimmung der Intelligenz muß unabhängig von der denkpsychologischen Analyse erfolgen. Denn das Wesen einer seelischen Disposition besteht nicht darin, daß sie bestimmte Akte vollziehe (z. B. solche des Kombinierens oder des Verstehens) oder daß sie bestimmte Phänomene (Gegenstandsgedanken, Beziehungsbewußtsein usw.) erzeuge, sondern, daß sie gewisse Teilzwecke des persönlichen Lebens verwirkliche. Wenn daher die von mir versuchte Intelligenzdefinition teleologischer Natur ist, so ist dies nicht ein Notbehelf, weil eine Feststellung der einzelnen bei der Intelligenz hervortretenden Denktätigkeiten zu umständlich oder noch nicht möglich wäre, sondern es ist eine aus der Grundüberzeugung des Personalismus sich ergebende notwendige Folgerung. Übrigens darf es gesagt werden, daß die folgende Begriffsbestimmung sich bereits in der wissenschaftlichen Arbeit der letzten Jahre gut bewährt hat und von anderen Forschern übernommen worden ist.

Die Definition lautet:

Intelligenz ist die allgemeine Fähigkeit eines Indi-

viduums, sein Denken bewußt auf neue Forderungen einzustellen; sie ist allgemeine geistige Anpassungsfähigkeit an neue Aufgaben und Bedingungen des Lebens.

Zur Erläuterung sei folgendes gesagt:

Im gewöhnlichen Sprachgebrauch — und zwar nicht nur des Alltagslebens, sondern auch der Forscher — hat das Wort I. lange Zeit eine sehr schwankende und fließende Bedeutung gehabt. Namentlich die Psychiater waren geneigt, den Ausdruck viel zu weit zu fassen. Sie verstanden darunter den Inbegriff der geistigen Leistungsgebiete überhaupt, also aller derjenigen psychischen Eigenschaften, die nicht affekt- oder willensmäßiger Natur sind. In diesem Sinn gehören dann Merkfähigkeit und Lernfähigkeit, Unterschiedsempfindlichkeit, Zuverlässigkeit der Aussage, Summe der Kenntnisse usw. mit gleichem Recht und Gewicht zur I. wie die Fähigkeiten des Auffassens und Kombinierens, des Urteilens und Schließens, des Kritisierens und Definierens, des Sich-Zurechtfindens und Sich-Zuhelfenwissens usw.; und eine Frage, die für uns eine der wichtigsten wird: inwiefern die zuerst genannten Eigenschaften mit der eigentlichen I. zusammenhängen, wird dann sinnlos. In den letzten Jahren ist man aber von dieser weiten und unscharfen Begriffsbestimmung zurückgekommen. Man erkannte, daß die I. nicht nur gegen die Gemüts- und Willensbeschaffenheit eines Individuums abzugrenzen ist, sondern daß ihr auch innerhalb der geistigen Funktionen ein deutlich abgesteckter Platz angewiesen werden müsse.

Dies geschieht nun durch die von uns versuchte Definition.

a) Das Merkmal der Allgemeinheit der Fähigkeit sondert die I. vom Talent, dessen Wesenszug gerade in der Beschränkung der Leistungsfähigkeit auf ein inhaltliches Teilgebiet besteht. „Intelligent" ist dagegen derjenige, der unter den verschiedensten Bedingungen und auf den verschiedensten Gebieten neuen Situationen geistig gewachsen ist und jeweilig die durch die Lage geforderten Denkoperationen in zweckmäßiger Weise einleitet und vollzieht. Ist das Talent eine materiale (z. B. auf Sprache, Musik, Mathematik usw. gerichtete), so ist die I. eine formale Leistungsfähigkeit.

Nun ist es freilich von manchen Psychologen bestritten worden, daß diese „formale Allgemeinheit" der geistigen Anpassungsfähigkeit überhaupt existiere. Es wird dann behauptet, die sogenannte I. zerfalle in eine Reihe von Einzeleigenschaften, deren jede für sich eine bestimmte Höhe haben könne; aber ein zentraler Faktor, eine

für das Individuum durchgängige Stärke und Färbung seiner Denkfähigkeit sei nicht vorhanden. Eine solche Behauptung wird nun nicht nur durch die gesamte empirische Intelligenzforschung der letzten Jahre widerlegt, die imstande war, diese allgemeine Fähigkeit bei jedem Einzelmenschen in ihrer Gesamthöhe zu bestimmen; ihr steht ebenso die natürliche Erfahrung des Lebens Schritt für Schritt entgegen, welche die Kennworte „klug" und „dumm" erfolgreich anwendet, um die allgemeine Höhe der geistigen Anpassungsfähigkeit — nicht ihre Beschränkung auf ein Teilgebiet — zu kennzeichnen; und sie wird endlich durch die philosophische Überzeugung überwunden, daß der Mensch nicht nur eine Musterkarte mosaikartig aneinandergereihter Eigenschaften ist, sondern eine einheitliche Persönlichkeit, die einheitliche Zwecke verfolgt und für diese Zielstrebigkeit einheitlich gerüstet und vorbereitet ist. — Daß aber die „Einheitlichkeit" der I.-Anlage nicht starre Einfachheit bedeutet, daß sie vielmehr eine große innere Mannigfaltigkeit und Beweglichkeit einschließt, wird in den weiteren Darlegungen zu zeigen sein.

b) Das Merkmal des Neuen, auf welches die Einstellung erfolgt, trennt die I. vom Gedächtnis, dessen teleologisches Grundmerkmal die Bewahrung und Verwertung schon dagewesener Bewußtseinsinhalte und die Wiederholung früher geübter psychischer Tätigkeit ist[1]). Gedächtnis und Intelligenz bilden so typische Vertretungen der beiden in jedem Menschen lebenden Zweckrichtungen auf Erhaltung und Entfaltung. Gedächtnis ist konservativ und rückwärts gewandt, I. progressiv und in die Zukunft weisend. Das Gedächtnis hat seine Rolle da zu spielen, wo der Mensch vor gleiche Situationen und Aufgaben gestellt ist, wie sie ihm früher begegnet sind; und es bewirkt, daß er auf diese bekannten Reize mit immer steigender Treffsicherheit und Kraftersparnis zu antworten vermag. Die I. tritt dort ins Spiel, wo die ständig wechselnde Konstellation der Umweltverhältnisse sich zu einem für das Individuum neuartigen Reiz formt, auf welchen auch nur mit einer für das Individuum neuartigen, momentan erst zu erzeugenden geistigen Tätigkeit reagiert werden kann.

c) Aber diese Neutat muß als Kennzeichen der I. auch nach der anderen Seite hin abgegrenzt werden. Sie ist nicht die eigent-

[1]) Gedächtnis ist hier in der weitesten Bedeutung, etwa im Sinne der „Mneme" zu fassen. Auch der Instinkt als Gattungsgedächtnis, Gewöhnung und Übung als psychophysisches Gedächtnis gehören dazu.

lich schöpferische Tat, die, ganz von innen quellend, erst der Umwelt neue Gestalten des Daseins aufzwingt; sondern eine Handlungsweise, die — auch in ihrer Neuartigkeit — durch die neue Konstellation der Welt da draußen bedingt ist. Darum gehört das Merkmal der Anpassung in die Begriffsbestimmung der I.; hiermit wird die Abhängigkeit der Leistung von den äußeren Momenten (der Aufgabestellung, den Forderungen des Lebens) hervorgehoben und die I. von der wahren Schaffensfähigkeit oder Genialität ebenso wie von der wahren Selbstherrlichkeit des Geistes, der Weisheit, geschieden.

Nur darf freilich der Begriff der Anpassungsfähigkeit nicht zu eng genommen, vor allem nicht nur auf Gegenwartssituationen beschränkt werden. Es gibt typische Äußerungen der I., die sich nicht in „Geistesgegenwart" erschöpfen, d. h. darin, daß auf die Forderung des Augenblicks die zweckmäßigste geistige Einstellung gefunden wird. Gerade die höchsten Bekundungen der I. sind die vorausschauenden; und die Unabhängigkeit von der Augenblickslage kann hier so groß werden, daß die Handlung einen durchaus spontanen, rein innerlich begründeten Eindruck macht. Dennoch gilt auch hier der Begriff der Anpassung — nur daß sich der Mensch nicht mehr an wirkliche, sondern an mögliche oder wahrscheinliche Umweltbedingungen anpaßt. Der intelligente Feldherr muß bei seinen Gegenwartsbefehlen mit den zu erwartenden Änderungen der Gefechtslage rechnen, der intelligente Börsenmann legt seinen heutigen Spekulationen die künftige Welternte oder eine in Aussicht stehende politische Verwicklung zugrunde — stets liegt eine vorwegnehmende Reaktion vor; aber ein bloßes Re-Agieren auf Faktoren, die dem Handelnden als äußerliche gegenüberstehen, und demnach ein Sich-Anpassen ist auch hier das Charakteristische des Verhaltens.

Diese für alle I. wesentliche Gebundenheit an die äußeren Lebensforderungen ist es, die auch ihrer moralischen Bewertung eine ganz bestimmte Färbung gibt. Die Nur-Intelligenz bietet das Bild einer geistigen Fähigkeit, deren Stärke in peinlichem Widerstreit zu dem geringen Maße von Innerlichkeit stehen kann. Die vielseitige Anpassungsfähigkeit an verschiedenste Umstände tritt leicht in einen gewissen Gegensatz gegen die Treue des Menschen zu sich selbst, gegen die spontane Auswirkung eines eigenen inhaltsbestimmten Persönlichkeitsideals und gegen die unbedingte Hingabe an überpersönliche Ideale.

Darum ist die I. an sich nicht als Wert schlechthin aufzufassen, vor allem nicht als richtunggebender Wert für bestimmte Kulturgebiete, wie es z. B. das Erziehungs- und Berufsleben ist; hier müssen wir uns vor einem falschen Intellektualismus hüten oder, so weit wir schon in ihm stecken, wieder aus ihm herauszukommen suchen. Wohl aber darf und soll die I. gewertet werden als Teilfähigkeit, die eingeordnet ist in den gesamten geistigen Kräftehaushalt des Menschen — wie ja auch die Anpassung wahrlich nicht der höchste, wohl aber ein notwendiger und an seiner Stelle wertvoller Teilzweck im gesamten Telos der Persönlichkeit ist. I. an sich ist nur Mittel, Rüstzeug; wertvoll (oder wertwidrig) wird sie erst durch die positiven Ziele, in deren Dienst sie verwandt wird.

Einer solchen Auffassung entspricht auch die Grundüberzeugung der personalistischen Psychologie. Nach ihr gibt es keine getrennten Seelenvermögen (wie es die alte Vermögenstheorie wollte), sondern alle Dispositionen sind nur Teilstrahlen der einheitlichen Zielstrebigkeit des Menschen, seines Erhaltungs- und Entfaltungsstrebens; diese Teilstrahlen kreuzen und verschmelzen sich in mannigfacher Weise, stehen in engerer oder loserer „Korrelation" miteinander. So sollte zwar unsere Definition die I. gedanklich scharf abgrenzen gegen andere Seiten der geistigen Persönlichkeit, und wir haben diese Abgrenzung gegen Talent, Gedächtnis, Genialität soeben vollzogen; aber es ist nur eine gedankliche, keine reale Trennung. Es gibt z. B. nie ein isoliertes Funktionieren der I. für sich und des Gedächtnisses für sich, vielmehr ist jede Gedächtnisleistung mehr oder minder von I.-Funktionen durchsetzt und umgekehrt; über das Mehr oder Minder dieser Zusammenhänge kann nur die Korrelation der geprüften Symptome Auskunft geben. Aber eben wegen dieser Zusammengesetztheit jedes wirklichen seelischen Vorgangs scheint mir obige Definition als regulatives Prinzip der weiteren Forschung erforderlich; irgend eine Wahrnehmungs-, Gedächtnis-, Aufmerksamkeits-, Schaffensleistung usw. ist zugleich eine I.-Leistung insofern, als sie eine geistige Neueinstellung auf neue Anforderungen enthält.

Es sei noch besonders hervorgehoben, daß diese enge Verknüpftheit der I. mit anderen Seelenfunktionen nicht nur auf das geistige Gebiet (Gedächtnis, Phantasie, Sinneswahrnehmung) beschränkt ist, sondern auch auf das Willens- und Gemütsleben hin-

übergreift. Es sind die praktischen Willenshandlungen vielfach intelligenzdurchsetzt (von dieser praktischen I. wird weiter unten zu sprechen sein), und es ist auch umgekehrt die intelligente Denkleistung in hohem Grade ein Willensakt und von Willenseigenschaften: Konzentration, Ausdauer, Zuverlässigkeit und Selbstzucht, abhängig.

Bezüglich des Gemütsanteils empfiehlt es sich vielleicht, zwischen Intelligenz und Intellektualität zu scheiden. Intelligenz ist die Fähigkeit zu geistigen Neuanpassungen, Intellektualität die Neigung hierzu. Beim „intelligenten" Menschen ist nur von dem geistigen Mittel die Rede, über das er zur Erreichung beliebiger Zwecke verfügt; beim „Intellektuellen" wird die Handhabung dieses Mittels selbst zu einem gefühlsbetonten Zweck; das Leben im Gedanklichen wird ihm zu einer Gemütssache. Beide Seiten sind meist eng miteinander verknüpft; doch gibt es auch Unstimmigkeiten und wirkliche Gegensätze zwischen ihnen, von denen später die Rede sein wird.

Alle eben angedeuteten Verschränkungen der I. mit anderen persönlichen Eigenschaften dürfen nicht ganz aus dem Auge gelassen werden, auch dort nicht, wo wir aus methodischen Gründen die I. isolierend behandeln werden, als ob sie ein selbständiger Faktor der Persönlichkeit wäre.

2. Andere Begriffsbestimmungen der Intelligenz.

Auf dem Gebiet der I.-Definition herrscht zurzeit noch eine große Verworrenheit, die so recht in die Augen fällt, wenn man die dankenswerte Zusammenstellung solcher Bestimmungsversuche durch Anschütz (*1*) und Höper (*24*) liest.

So ist schon darüber keine Einheitlichkeit vorhanden, ob man unter I. eine bei allen Menschen vorhandene Fähigkeit verstehen, oder ob man den Begriff auf die nur gewissen Menschen zukommende „hohe Intelligenz" beschränken soll. So sagt z. B. Meumann (*10*, S. 9): „Mit dem Begriff der Intelligenz bezeichnet der Sprachgebrauch ein höheres Gesamtniveau des intellektuellen Lebens oder einen qualitativ höheren Gesamttypus der Begabung." Sollte man diese Bedeutung des unwissenschaftlichen Sprachgebrauchs in die Psychologie übernehmen, so würde dies etwa einer Physik gleichen, die in der „Wärmelehre" nur die Temperaturen über 0 Grad (also die von Laien als „Wärme" bezeichneten Erschei-

nungen) behandeln wollte. Für den Psychologen ist I. eine generell menschliche Disposition, die sich nun freilich nach Grad und Art stark differenziert; aber alle diese Differenzierungen, Schwachsinn ebenso wie Hochbegabung, fallen doch unter den gemeinsamen Oberbegriff der Intelligenz.

In Lehrerkreisen wird zuweilen versucht, den Ausdruck „Intelligenz", da er ein Fremdwort sei, zu beseitigen und durch „Begabung" oder „Fähigkeit" zu ersetzen. Der Kampf gegen überflüssige Fremdwörter ist nun gewiß voll berechtigt; aber an falscher Stelle eingesetzt, kann die Verdeutschung nur zur Begriffsverwirrung und -verwischung führen. Die zutreffende Verdeutschung für I. wäre nicht „Begabung" schlechthin, sondern „geistige Allgemeinbegabung", um die I. von den Willensbegabungen und von den geistigen Sonderbegabungen (Talenten) zu unterscheiden. Aber der Gebrauch jener langen Wortbildung wäre so unbequem, daß man hier ruhig das wissenschaftliche Heimatsrecht des Fremdwortes anerkennen sollte.

Sachlich dreht sich der Hauptstreit um die Frage, ob wir überhaupt von einer einheitlichen Gesamtfähigkeit der „Intelligenz" sprechen dürfen, oder ob das Wort nur ein (mehr oder minder gerechtfertigter) Sammelname für eine Reihe disparater Funktionen sei. Als erster hat Spearman (*35—37*) die Frage in die Erörterung geworfen und das Vorhandensein einer „general ability" behauptet. Er begründete diese These durch Korrelationsuntersuchungen; d. h. er wies nach, daß eine Reihe sehr verschiedener experimentell prüfbarer Leistungsformen an den geprüften Individuen gemeinschaftlich variieren; ein Individuum, das in einigen geprüften Gebieten Gutes leistete, bot die Wahrscheinlichkeit, auch in anderen Gutes zu leisten und umgekehrt. Es muß also, so schloß er, einen „gemeinsamen Zentralfaktor" geben, der seine Wirkung auf sehr verschiedene Inhalte des Seelenlebens erstreckt. Das nähere Wesen dieses Zentralfaktors bezeichnet er mit großer Vorsicht als eine allgemeine „plastische Funktion des Zentralnervensystems", die das eine Individuum gegenüber dem anderen in gesteigertem Maße besitzt.

Die Korrelation der Einzelfunktionen hat sich auch weiterhin als starke Stütze der Auffassung erwiesen, daß es für jeden Menschen einen allgemeinen geistigen Habitus gibt, der verschiedenen Teilgebieten eine gemeinsame Färbung und Stärke verleiht. Allerdings liegen die Verhältnisse nicht so einfach, wie Spearman glaubt; gewisse Teilgebiete (wie z. B. die von Sp. noch stark in den Vordergrund geschobene Unterschiedsempfindlichkeit) stehen nur in loserem Zusammenhang, andere (wie insbesondere eigent-

liche Denk- und Urteilsleistungen) in engerem Zusammenhang zur allgemeinen I. Auch von Mensch zu Mensch ist die Art, wie sich die „general ability" in den einzelnen Leistungsgebieten bekundet, noch verschieden. Dennoch ist die Existenz einer solchen gemeinsamen Geisteshaltung, insbesondere wenn man sie in diesem beweglichen Sinne als bloße Rahmenform der geistigen Einzelleistungen auffaßt, eines der sichersten Ergebnisse der bisherigen Forschung auf unserem Gebiete.

Recht im Gegensatz hierzu bestreitet Ziehen (*45*) durchaus die Einheitlichkeit der I. Unter den Leitsätzen, die er neuerdings über das Wesen der Beanlagung aufgestellt hat, befinden sich die folgenden: „Die intellektuelle Beanlagung zerfällt in zahlreiche Anlagen, die untereinander in verwickelten Zusammenhängen stehen. Eine allgemeine Verstandesfähigkeit oder I. existiert nicht." „Es ist daher auch nicht zulässig, von einer Gesamtbegabung zu sprechen oder gar diese hypothetische Gesamtbegabung zahlenmäßig auszudrücken." „Unter der Bezeichnung ‚Intelligenz' faßt man zweckmäßig einige ganz bestimmte intellektuelle Anlagen zusammen, nämlich das Gedächtnis und seine mannigfachen Unterarten, die Begriffsbildung und die sogenannte Kombination." Diese Anschauung entspricht der Gesamtüberzeugung Ziehens von dem bloß assoziativen, mosaikmäßigen Aufbau des Seelenlebens[1]). Die Hineinnahme des Gedächtnisses in den I.-Begriff verhindert die wichtige Scheidung zwischen den besonderen Aufgaben und Zielen von I. und Gedächtnis, die wir oben vornahmen; die Aufzählung der zur I. gehörigen Denkleistungen ist dagegen ganz unvollständig, denn Erkennen, Verstehen, Urteilen, Kritisieren, Zerlegen sind ebenso wichtige Äußerungsformen der I. wie Begriffsbildung und Kombination.

Einen ähnlichen Standpunkt wie Ziehen nimmt auch Höper (*24*) ein, indem er erklärt: „Die Auffassung einer ‚Intelligenz' in Form einer einheitlichen psychischen Funktion wird abgelehnt." Trotzdem versucht er eine Begriffsbestimmung, von der er freilich die sonderbare Meinung hegt, daß sie „den Begriff als solchen aufhebt". Sie wird weiter unten (S. 11) angeführt.

Ausführlich beschäftigt sich Meumann (*9*) mit dem Begriff

[1]) Freilich spricht Ziehen in einem weiteren Leitsatz von der „Unterscheidung minderbegabter, mittelbegabter und überbegabter Individuen", die er für gerechtfertigt hält. Hiermit wird also die „Gesamtbegabung" und deren gradweise Variation stillschweigend anerkannt, die vorher ausdrücklich abgelehnt worden war

der I. Er erkennt an, daß wir das Recht haben, von einer „einheitlichen Gesamtverfassung des Individuums" zu sprechen; sie beruht auf der einheitlichen Gestaltung des gesamten Seelenlebens, die von einer einzigen zentralen Fähigkeit ausgeht (S. 705). Er gibt ferner zu, daß man den teleologischen Gesichtspunkt heranziehen müsse, um das Wesen der I. ganz zu verstehen, hält aber eine Begriffsbestimmung, die allein teleologisch gerichtet ist, für unzureichend. Es müsse vielmehr zugleich auch die Fülle der psychologischen Tatbestände mit berücksichtigt werden. Nun ist diese Mitberücksichtigung selbstverständlich, und sie wird auch des weiteren in diesem Buche erfolgen; wie aber diese Vielheit der bei der I. beteiligten Seelenprozesse in die Begriffsbestimmung mit aufgenommen werden kann, bleibt auch trotz der ausführlichen Darlegung M.s unverständlich. Denn er gibt zwar an verschiedenen Stellen Schilderungen, aus denen hervorgeht, daß die I. über das bloße Aufnehmen und Reproduzieren hinausgeht und Selbständigkeit sowie schöpferische Fähigkeit des Denkens oder der Phantasie voraussetzt; nirgends aber versucht er, daraus eine rein im Psychologischen bleibende Definition der I. abzuleiten. Die Hineinbeziehung des „Schöpferischen" bildet aber einen weiteren Nachteil der M.schen Bestimmungen; dadurch verliert nämlich der Begriff der I. nach oben hin seine Schärfe, sowie er sie bei Ziehen nach unten hin (zum Gedächtnis) verloren hatte: die Grenze zwischen der bloßen geistigen Anpassungsfähigkeit und der eigentlichen Produktivität ist bei M. völlig verwischt[1]).

Aus den verschiedenen Beschreibungen, die Meumann von der I. gibt, greift Anschütz einen einzelnen Tatbestand heraus: er sieht das Wesen der I. darin, Beziehungen aufzufinden und herzustellen. Damit knüpft er an einen der frühesten Versuche der Begriffsbestimmung an, nämlich an Ebbinghaus' Gleichsetzung von geistiger Leistungsfähigkeit mit „Kombinationsfähigkeit" (1897) (*60*). Hiergegen spricht erstens der Umstand, daß das „Beziehen" nur eine von den mannigfachen Betätigungsweisen der I. ist und daher nicht den Begriff erschöpfen kann, zweitens die Tatsache, daß nicht das Herstellen irgend welcher, sondern das Herstellen sinnvoller, von der Situation geforderter Beziehungen kennzeichnend für die I. ist. (Beziehungen stellt ja

[1]) Diese Schwächen der Meumannschen Begriffsbestimmung werden in treffender Weise von Hasserodt (*23*) charakterisiert, der im wesentlichen meiner Definition zustimmt.

auch der Ideenflüchtige her!) Es ist eben nicht möglich, die I. ohne Berücksichtigung eines teleologischen Gesichtspunktes zu definieren. Dies berücksichtigte Höper, der deshalb die Anschützsche Bestimmung folgendermaßen ausbaut: „I. nenne ich die Fähigkeit des Individuums, Beziehungen aufzufinden und herzustellen. Diese Fähigkeit ist orientiert an den Sachgebieten des Lebens. Sie kann sich äußern in mehr produktiver oder mehr rezeptiver Weise, mehr analysierend oder mehr synthetisch" (s. oben, S. 9). Eine Definition der I., die der meinigen sehr ähnelt, gibt neuerdings Claparède (*20*, *21*). Sie ist ihm „die Fähigkeit, vermittels des Denkens neue Probleme zu lösen"; „ein Werkzeug der Anpassung, das ins Spiel tritt, wenn die anderen Anpassungsmittel, Instinkt und Gewöhnung, versagen." Kennzeichnend für jeden Intelligenzakt sind ihm drei Hauptoperationen: das Fragen, das Suchen der Lösung (Aufstellen einer Hypothese), die Erprobung (Verifikation der Hypothese).

II. Kapitel.

Die Bedingungen der Intelligenz.

Wir bezeichneten die I. als eine dem Menschen innewohnende Disposition; und so gelten auch alle Bestimmungen, welche die personalistische Psychologie für die Dispositionen zu treffen sucht[1]), auch für die Intelligenz. Dispositionen sind dauernde potentielle Ursächlichkeiten für die Einzelakte und Bewußtseinsinhalte des Menschen; sie sind aber als solche ergänzungsbedürftig und mehrdeutig; d. h. sie können aus sich heraus allein nicht das Tun des Menschen bestimmen, sondern bedürfen des Zusammenwirkens (der „Konvergenz") mit äußeren Einflußfaktoren, um zu realer psychischer Wirksamkeit zu führen. Bei jeder Disposition, so auch bei der I., ist daher das Verhältnis der inneren Bedingtheit, die im Wesen der Disposition selbst liegt, und der äußeren Bedingtheiten, die als „Umwelt" (Erziehung und Unterricht, Lektüre, Vorbild, Spiel usw.) mitwirken, ein Grundproblem.

Die Unvollständigkeit und Vieldeutigkeit, die jeder Disposition als solcher anhaftet, führt zu einer Zweiteilung der Dispositionen in „Anlagen" und „Eigenschaften"; und es ist von entscheidender Wichtigkeit, daß der Begriff I. in beiden Bedeutungen angewandt

[1]) Vgl. Die menschliche Persönlichkeit (*16*), Kap. II, Die Dispositionen.

werden kann. Die I. ist zweifellos eine angeborene Disposition; angeboren aber ist sie nur als Anlage, d. h. als noch nicht fest abgezirkelte Betätigungsmöglichkeit mit breitem Spielraum, als eine Vieldeutigkeit, die erst der Vereindeutigung im Laufe der Entwicklung und unter Beteiligung der Einflüsse des Lebens harrt. Je mehr diese Vereindeutigung fortschreitet, je mehr sich feste Betätigungsweisen herausbilden, um so mehr wird aus der Anlage eine Eigenschaft, die nun nach Art und Grad ein gemeinsames Erzeugnis der inneren und äußeren Bedingungen ist. Durch diesen Sachverhalt wird die später zu besprechende Aufgabe der I.-Untersuchung außerordentlich erschwert, zugleich aber eines der wichtigsten Probleme berührt, das durch eben diese Untersuchung geklärt werden soll; denn was wir unmittelbar durch irgend eine Methode der I.-Untersuchung feststellen, sind Äußerungen der I. als einer Eigenschaft, wie sie augenblicklich als Konvergenzerzeugnis von Anlage und Umwelt beschaffen ist; und so wichtig auch die Feststellung dieser Gegenwartseigenschaft sein mag, so muß es doch unser Ziel sein, aus ihr rückwärts die etwa zugrunde liegende ursprüngliche (d. h. angeborene) I.-Anlage und vorwärts die etwa aus dieser Anlage zu erwartende Zukunftsentwicklung abzuleiten. Zugleich kann eine solche Analyse dann dazu führen, den Anteil auch der äußeren Faktoren an dem Zustandekommen der festgestellten I.-Beschaffenheit zu bestimmen.

Sehen wir genauer zu, so zerfällt jede der beiden Gruppen von Bedingungen, aus deren Konvergenz die tatsächliche I. hervorgeht, in zwei Untergruppen.

Die inneren Bedingungen der I. sind die eigentliche Anlage und das Alter. Von der Anlage sprachen wir schon; sie setzt trotz ihrer Vieldeutigkeit doch von vornherein gewisse Grenzen, innerhalb deren allein sich die Entwicklung — auch bei noch so günstigen Umwelteinflüssen — vollziehen kann; und sie enthält auch von vornherein wenigstens die allgemeine Richtung, nach der sich die I. entfalten wird. Selbst die beste Schule wird aus einem Dummkopf keinen klugen Menschen machen; auch die einseitigste Beeinflussung wird aus einer der Anlage nach analytischen I.-Richtung keine vornehmlich synthetische Geistesfähigkeit erzeugen. Diese Anlage ist also keine Prädestination, wohl aber eine Prädisposition; wir sollen nicht mit ihr wie mit einer starren Größe fatalistisch rechnen, sondern sollen auf sie einzuwirken suchen, indem wir von ihrer Bildsamkeit und Schmiegsamkeit Gebrauch

machen; aber wir können sie nur heranbilden, entwickeln und veredeln, wenn wir die in ihr selbst gegebenen Zielstrebigkeiten anerkennen und beachten.

Suchen wir die I.-Anlage selbst noch weiter kausal zurückzuführen, so stoßen wir auf die Faktoren der Vererbung und des Geschlechts. Daß bei der I. die Erblichkeit eine Rolle spielt, ist unbestritten (wenn auch noch nicht in allen Einzelheiten klargestellt); Alltagserfahrung und psychologische Massenerhebungen stimmen darin überein, daß Kinder intelligenter Eltern die größere Wahrscheinlichkeit haben, ihrerseits intelligent zu werden, als Kinder unintelligenter Eltern, und umgekehrt.

Eine massenstatistische Untersuchung von W. Peter (*33*) bearbeitet die Schulzeugnisse dreier Generationen aus ländlichen Gegenden Württembergs. Es ergab sich, daß die Zeugnisse der Kinder im Durchschnitt um so schlechter waren, je schlechter die Zeugnisse der Eltern waren. (Allerdings wichen die Durchschnittswerte der Kinder nicht so stark voneinander ab, wie die Zeugnisnoten der Eltern; Galtons Gesetz des erblichen Rückschlags.) Das gleiche Verhältnis bestand zwischen den Zeugnisnoten der Großeltern und der Enkel.

Eine zweite Massenerhebung ist von Heymans und Wiersma (*25*) angestellt worden; hier wurden Ärzte gebeten, über Eltern und Kinder ihnen genau bekannter Familien Angaben über die seelischen Haupteigenschaften zu machen. Die statistischen Berechnungen ergaben, daß gerade bei den I.-Eigenschaften die Übereinstimmungen zwischen der älteren und der jüngeren Generation besonders deutlich waren; und zwar handelte es sich meist um gleichgeschlechtliche Erblichkeit: die Söhne waren den Vätern, die Töchter den Müttern ähnlich.

Dies Ergebnis führt zu der Bedeutung des Geschlechts für die I. Eine solche Bedeutung ist fraglos vorhanden; aber sie besteht nicht, wie es Möbius (*32*) und andere früher glaubten, in einer Minderwertigkeit der weiblichen Intelligenz, sondern in einer Andersartigkeit; und diese scheint gar nicht auf primär intellektuellen Gründen zu beruhen, sondern in Gemüts- und Willensanlagen zu wurzeln, die dann auf das geistige Gebiet übergreifen. Für die I. der Frau dürfte wohl im allgemeinen gelten, was ich über die des heranwachsenden weiblichen Geschlechts an anderer Stelle (*41*) ausführte: „Das Entscheidende ist nicht, daß die Mädchen eine geringere Fähigkeit, sondern daß sie ein geringeres Bedürfnis nach spontanem intellektuellem Tun haben; nicht ihre Intelligenz, sondern ihre Intellektualität steht hinter der des Knaben zurück. Diese geringere Neigung wurzelt einerseits in der Rezeptivität, dem schnelleren Zufriedensein mit dem Gegebenen, andererseits in letzten Einstellungen des Interesses, die dem Abstrakten, Unpersönlichen, Gefühlsfremden — und damit dem spezifisch

Logischen — entgegenwirken." Hiermit steht es nicht in Widerspruch, daß nach den umfassenden Materialien zur psychischen Geschlechtsvergleichung, die Lipmann (28) zusammengestellt hat, die Neigung zu rezeptiver geistiger Betätigung (also zum Lernen, Studieren, Lesen) beim weiblichen Geschlecht eine größere Stärke und Verbreitung besitzt als beim männlichen, ebensowenig, daß bei I.-Prüfungen, die ja nur die reaktive I. feststellen, keine eindeutigen Unterschiede zugunsten des einen oder anderen Geschlechts hervorgetreten sind.

Der Umstand, daß der Unterschied zwischen den Geschlechtern gerade in den spontanen Seiten der Persönlichkeit wurzelt, spricht gegen die öfter zu hörende Annahme, daß nur die Verschiedenheit der Erziehung und Schulbildung, also äußere Faktoren, als Ursachen in Betracht kommen; er macht es vielmehr höchst wahrscheinlich, daß es sich hier wirklich um innerliche Anlagedifferenzen handelt.

Man darf nun übrigens nicht glauben, daß aus Vererbung und Geschlecht die I.-Anlage des einzelnen Individuums mit mathematischer Eindeutigkeit abgeleitet werden könnte. Auch sie geben nur Möglichkeiten und Wahrscheinlichkeiten; stets aber bleibt ein irrationaler Rest. Jedes Individuum ist nach seiner Intelligenzanlage nicht nur seinen Eltern ähnlich und seinem Geschlecht entsprechend, sondern besitzt eine letzte singuläre und unableitbare Eigenart der Anlage, durch die sogar gelegentlich die Einflüsse der beiden genannten Faktoren überschattet werden können. Es ist vielleicht nicht unnötig, insbesondere gegenüber einer beliebten Überwertung des Vererbungsgrundsatzes, ausdrücklich auf dieses individuelle Anlagemysterium hinzuweisen, das jeder Vorhersage oder gar Berechnung spottet.

Die allmähliche Entwicklung der I.-Anlage zur Eigenschaft ist nun zunächst noch von einem weiteren inneren Faktor abhängig, dem Alter. Es lebt im Individuum eine eingeborene Gesetzlichkeit des Reifens, auf Grund deren die allgemeine Anlage eine Fähigkeit nach der anderen zur Auswirkung bringt. Wohl ist das Tempo dieser Entwicklung von äußeren Hilfen oder Hemmungen mitbestimmt; aber diesen Einflüssen sind doch ziemlich enge Grenzen gesetzt, und die Reihenfolge, in der sich die Einzelfähigkeiten natürlicherweise herausstellen, ist ganz wesentlich innenbedingt. Hier erwächst der Forschung wieder eine neue Aufgabe

in den konkreten I.-Eigenschaften einer bestimmten Altersstufe zu scheiden, was auf die Stärke und Art der ursprünglichen Anlage, was auf die augenblickliche Entwicklungsphase zurückzuführen ist. Es gibt I.-Leistungen, die viel mehr vom Alter als von dem ursprünglichen Begabungsfonds abhängig sind, so daß schwachsinnige ältere in ihnen höher stehen als hochbegabte jüngere Individuen; es gibt andere, die fast unabhängig vom Alter allein aus der Beschaffenheit der Anlage zu erklären sind. Verwandt hiermit ist die Frage, ob irgend eine — positive oder negative — Abweichung der I. von der Norm der Altersstufe auf einer bloßen Beschleunigung bzw. Verlangsamung der Entwicklung, also auf Frühreife bzw. Spätreife beruht (ohne daß daraus die schließliche Endhöhe der Entwicklung zu entnehmen wäre), oder ob es sich um eine stärkere bzw. schwächere Grundbeschaffenheit der Anlage selbst handelt, die auch für das Endergebnis der Entwicklung eine Voraussage gestattet.

Mit diesen inneren Bedingungen der Anlage und des Alters konvergieren nun die äußeren Faktoren, die sich ihrerseits wieder in Unterrichts- und Milieueinflüsse gliedern.

Die Bedeutung der Schule für die Entwicklung der I. bedarf einer genaueren Analyse. Der Schulunterricht vermittelt materiale Wissensinhalte („Kenntnisse") und formale Übungen in der Verwertung dieser Kenntnisse, im denkenden Verarbeiten gegebener Stoffe, in der sprachlichen Umsetzung von Gedanken. Solange man nun das Gedächtnis und seine Leistungen einfach miteinbezieht in den Begriff der I. — wozu manche Forscher neigen —, ist es verständlich, daß man den Besitz gedächtnismäßig angeeigneten Schulwissens als ein Kriterium der I. betrachtet und deshalb Kenntnisprüfungen als I.-Prüfungen benutzt. Fassen wir aber I. im Sinne unserer Definition, so ist der bloße Besitz von Kenntnissen als solcher, wie er durch die Schule vermittelt wird, niemals ein Beweis für das Vorhandensein eines bestimmten I.-Grades, sondern höchstens ein Mittel, das die Entwicklung der I. erleichtert — zuweilen aber auch, wo die Kenntnisse nichts als unverdauter Gedächtnisballast blieben, eine Hemmung für die freiere und selbsttätige Denktätigkeit. So ist also der Einfluß der Schule auf die I. nach der materialen Seite keineswegs eindeutig.

Wie steht es mit der formalen Seite? Unverkennbar liegen hier positive Einflüsse vor; denn ein ganz großer Teil des Unter-

richts ist ja darauf angelegt, „das Denken zu üben". Schärfe der Begriffsbildung und Klarheit des Ausdrucks werden geschult, in die Abfolge der Gedanken wird System und Zusammenhang gebracht; es werden komplexe Gebilde zerlegt, Fehler gefunden und besprochen, Probleme formuliert. All dies hat formal-bildenden Wert; d. h. es wirkt über das Gebiet der ursprünglichen Übungsgegenstände hinaus und stärkt allgemeine geistige Kräfte, die sich nun auch auf anderen Gebieten mit gleicher Gewandtheit betätigen können. So ist I. als fertige Eigenschaft, oder kürzer: als „Fertigkeit" zweifellos von Umfang und Art des Schulunterrichts abhängig; und es ist im empirischen Einzelfalle gar nicht mehr scharf zu scheiden, wieviel an den aufgewiesenen I.-Leistungen auf diese erworbene Fertigkeit und wieviel auf die ursprünglich angelegte Fähigkeit zu schieben sei. Immerhin darf man nicht vergessen, daß auch die von der Schule gebotene intellektuelle Übung immer nur innerhalb der Spielraumbreite wirksam sein kann, die nun einmal durch die Anlage gewährt wird; und so schimmert doch durch alle schulbestimmten Leistungen immer die Grundfarbe der ursprünglichen und immanenten Begabung hindurch.

Auch sind es nur bestimmte Seiten der I., die durch die Schule besonders gepflegt werden. Die Denkübung vollzieht sich ganz vorwiegend an sprachlich geformten Stoffen; die in stummer Tat sich bekundende I. findet wenig oder keine Nahrung. Vor allem aber: die Schul-I. ist zu einem großen Teil reaktiver Natur; geschult wird die geistige Bewältigung von Aufgaben, die unmittelbar von außen an den Schüler herantreten, das Lösen von vorgelegten Problemen, das Verstehen von gebotenen Stoffen, das Beantworten von gestellten Fragen. Die spontane I. ist innerhalb der Schule viel schwerer erkennbar und viel weniger beeinflußbar. In ihr wird daher der Anlagefaktor in viel größerer Reinheit hervortreten.

Neben dem Schulunterricht, der mit Absicht die I. beeinflußt, stehen die zahlreichen „geheimen Miterzieher", die in ihrer Gesamtheit in dem unübersetzbaren Fremdwort „Milieu" zusammengefaßt werden[1]). Die soziale Lage der Eltern und die seelische Atmosphäre des Heims, Freundschaft und Verkehr, Lektüre und Reisen, die Reizungen und Ablenkungen der Großstadt und die

[1]) „Umwelt" ist der weitere Begriff; denn er umfaßt sowohl die absichtsvollen Einflüsse von Erziehung und Unterricht, wie die absichtsfremden Einwirkungen des „Milieus".

einfachen großzügigen Natureindrücke des Landes, Kino und Theater, die Erlebnisse der Straße und die Ereignisse der Politik —, sie alle formen mit am Menschen und so auch an seiner Geistigkeit. Sie sind imstande, die selbständige Denktätigkeit des Kindes zu fördern oder zu hemmen, in äußersten Fällen zu überreizen oder einzuschläfern; sie können von außen her Erscheinungen der Frühreife oder Spätreife bedingen, die von den innerlich begründeten gleichnamigen Entwicklungserscheinungen nicht leicht zu scheiden sind.

Im ganzen glaube ich, daß den „Milieu"-Einwirkungen ein noch stärkerer Anteil an dem Zustandekommen der I. beizumessen ist als den Schuleinflüssen. Während diese auf bestimmte Jahre und innerhalb dieser auf gewisse abgegrenzte Zeiten und Orte beschränkt sind, haben die Milieueinflüsse gleichsam Allgegenwart. Sie setzen mit dem ersten Lebenstage ein und können in den ersten sechs Lebensjahren, ehe noch die Schule beginnt, der I. schon ein ganz bestimmtes Gepräge gegeben haben. Sie laufen neben der Schule einher und dauern ununterbrochen an, wenn die Schule längst zu wirken aufgehört hat.

Aber alle diese Umwelteinflüsse — das muß das A und Z unserer Erörterung bleiben — „machen" nicht die I. Sie mögen noch so stark sein — mehr als den sichtbaren Oberbau zu liefern zu dem Fundament, das unsichtbar aber fest in der angeborenen Anlage verankert ist, vermögen auch sie nicht. Die I. als Fertigkeit ist stets nur eine Verwirklichung, eine Vereindeutigung der I. als Fähigkeit. Das eben ist der Grundgedanke jener „Konvergenz"-Theorie, die ebensowenig eine einseitige Erblichkeitslehre wie eine einseitige Milieulehre anerkennen kann[1]).

Für alle weiteren Behandlungen des I.-Problems aber ergibt sich, daß man überall, wo man von I. spricht, sich darüber klar sein muß, ob man sie im Sinne der Fertigkeit (Eigenschaft) oder der Fähigkeit (Anlage) meint. Unterläßt man dies, so wird gerade in den wichtigsten Fragen nie eine Klärung zu erzielen sein.

---

[1]) Über die Grundsätze der Konvergenzlehre vgl. Die menschliche Persönlichkeit (*16*), Kap. III.

III. Kapitel.

Die Verschiedenheiten der Intelligenz.

1. Intelligenzarten und -typen.

Die „Allgemeinheit" und „Einheitlichkeit" der I.-Funktion, von der wir oben sprachen, bedeutet nicht Einförmigkeit. Vielmehr kann die I. von Mensch zu Mensch nicht nur dem Grade nach, sondern auch der Art nach verschieden sein; und sie kann in demselben Menschen sich bald in der einen, bald in der anderen Art betätigen. Nur daß diese verschiedenen Artungen sich nicht in der Beschränkung auf bestimmte Inhaltsgebiete zeigen — das wären ja „Talente" —, sondern in der Verschiedenheit der formalen Verhältnisse, in denen die geistige Anpassungstätigkeit auftritt.

Sofern die Qualitäten der I. wechselnd im gleichen Individuum vorkommen können, sprechen wir von „I.-Arten"; sofern ihr Vorwiegen der I.-Disposition eines Menschen eine dauernde Färbung gibt, von „I.-Typen".

Die erste hierher gehörige Scheidung haben wir schon früher (S. 13) angedeutet: es ist die zwischen reaktiver und spontaner I.[1]).

Im letzten Grunde ist ja jede I. reaktiv, denn die Anpassung gehört zu ihrem Begriff; aber es ist ein großer Unterschied, ob diese Reaktion nur auf gegebene Gegenwartsreize oder auf mögliche Zukunftsreize erfolgt. Es gibt Intelligenzen, die einen ziemlich hohen Grad haben mögen, aber mit einer inneren Trägheit behaftet sind, so daß sie des jedesmaligen Anstoßes von außen bedürfen, um sich zu regen. Sie warten ab, lassen die Anforderungen und Aufgaben an sich herankommen und verstehen dann, sich trefflich mit ihnen auseinanderzusetzen und abzufinden. Es sind diejenigen, die in allen von außen weitgehend geregelten Lebensverhältnissen, in der Schule, in eng umschriebenen Berufstätigkeiten ihr Bestes geben. Dann aber gibt es die „spontanen" Intelligenzen, die auch zugleich eine lebhafte Intellektualität haben. Sie warten nicht, bis ihre geistige Arbeit herausgefordert wird, sondern neigen dazu, sie von sich aus ins Spiel zu setzen. Sie erleben Probleme, leiden unter ihnen und suchen sich mit ihnen auseinanderzusetzen; sie nehmen künftige Situationen vorweg, entwerfen Pläne, treffen vorschauende

[1]) Grundsätzliches über Reaktionen und Spontanaktionen und ihre Beziehungen findet sich in „Menschliche Persönlichkeit", Kap. IV.

Maßnahmen; und auch da, wo zunächst durch äußeren Anstoß eine Reaktion von ihnen gefordert wird, bleiben sie nicht bei der bloßen Antwortshandlung stehen, sondern bauen den eingeleiteten Denkvorgang weiter aus, stellen Fragen, die über den augenblicklichen Stand der Angelegenheit hinausgehen usw.

Natürlich gibt es zwischen der ausgesprochen reaktiven I. und der ausgesprochenen Spontaneität alle denkbaren Übergänge; und auch nach oben hin sind zwischen der spontanen I. und dem eigentlich schöpferischen Verhalten die Grenzen flüssig. Dennoch scheint es berechtigt, hier von typischen Unterschieden zu sprechen; denn das Vorwiegen der Tendenz zu bloßer Antworthandlung oder zu innenbedingter intelligenter Selbsttätigkeit ist doch bei einer großen Zahl von Menschen eindeutig festzustellen.

Der spontan Intelligente wird sich vor allem dort bewähren, wo freierer Selbsttätigkeit Spielraum gelassen ist. Er wird in gebundenen Situationen mit streng umgrenzten Aufgaben oft weniger leisten als der reaktive, dafür an führenden Stellen, in denen Initiative und Selbstverantwortlichkeit gefordert wird, oft überraschenderweise zeigen, was er vermag. Darum wäre nichts verhängnisvoller, als wenn man die Leistungsfähigkeit in reaktiver I. zum einzigen Maßstabe für den Aufstieg zu führenden Stellen ansehen wollte!

Eine andere Typenscheidung, die mit der eben behandelten nicht verwechselt werden darf, ist die zwischen objektiver und **subjektiver I.**[1]).

Daß Objektivität und Subjektivität zwei grundverschiedene Typen der Stellungnahme sind, die ihren Geltungsbereich über eine große Anzahl seelischer Tätigkeiten ausdehnen, ist an anderer Stelle von mir ausgeführt worden; hier handelt es sich darum, wie sie innerhalb der Intelligenz zum Ausdruck kommen. Geistige Anpassung ist stets die Herstellung einer zweckmäßigen Beziehung zwischen Subjekt und Objekt; sie kann aber entweder dadurch erzielt werden, daß das Subjekt die Tatsächlichkeiten und Gesetzmäßigkeiten der Objekte zur Richtschnur des eigenen Handelns macht, oder dadurch, daß das Subjekt an dem Objekt diejenigen Handlungen, Auslesen und Deutungen vornimmt, die den inneren

[1]) Vgl. hierzu: Diff. Psychol. (*12*), S. 210ff. und: Über I.-Stadien und -Typen beim Aussageversuch (*42*), S. 300ff.

Bedürfnissen des Subjekts selbst entsprechen. Von objektiver bzw. subjektiver Intelligenz kann man freilich nur dort sprechen, wo die Zweckmäßigkeit der Beziehung von Objekt und Subjekt bei diesen Handlungsweisen gewahrt bleibt. Darum ist es für uns wichtig, daß sowohl Objektivität wie Subjektivität sehr Verschiedenes und sehr Verschiedenwertiges bedeuten kann. Eine geringe Beteiligung des subjektiven Einschlags, also „Objektivität", kann einerseits auf träger Gleichgültigkeit, auf einem passiven Sichschiebenlassen vom Objekt beruhen, andererseits auf einer straffen Selbstzucht gegenüber subjektiven Eingebungen und Launen, auf der Ehrfurcht vor den Tatsachen, auf jener Überzeugung, daß man die Natur am besten beherrscht, indem man sich ihren Gesetzen fügt. Und entsprechend bei der Subjektivität: sie kann einerseits auf Hemmungslosigkeit beruhen, auf dem Unvermögen, individuelle Laune und beliebige Vorstellungsverkettung durch objektive Forderungen zu normieren, andererseits auf starker innerer Aktivität, die das Äußere bewältigt und sich zu eigen macht, indem sie es mit der eigenen Farbe durchtränkt, ihm die besonderen Seiten abgewinnt, die der Besonderheit des Ich entsprechen. Nur die beiden jeweils an zweiter Stelle genannten Verhaltungsweisen können objektive bzw. subjektive Intelligenz heißen.

Eine Reihe von experimentellen Befunden, die den Unterschied des objektiven und subjektiven Typs nachweisen, sind in der Differentiellen Psychologie aufgezählt worden.

An Erkennungsaufgaben bei tachistoskopischer Darbietung hat Messmer (*30*), an psychophysischen Urteilen über allmählich sich ändernde Reize habe ich selbst (*39*) jene Typengliederung exakt untersucht. Ein besonderes ergiebiges Gebiet stellen Aussage- und Bildbeschreibungsversuche dar; unter diesen ist die an erwachsenen Prüflingen angestellte Untersuchung von Baerwald (*19*) die psychologisch bedeutendste[1]). Es handelt sich um ein den Beschauern unbekanntes und nicht leicht deutbares Bild, über das schriftliche Darstellung verlangt wurde. B. findet hier die deutliche Scheidung eines „beschreibenden" und eines „selbständigen" Typs. Die Charakteristik des selbständigen Typs besteht, wie zahlenmäßig erwiesen wird, in folgenden Merkmalen: Neigung, den Zusammenhang des Ganzen zu beobachten, Vergleichungen vorzunehmen, über nicht gleich Verständliches Konjekturen aufzustellen, Kritik zu üben, endlich die Beziehungen zur eigenen Person stark zu betonen (Baerwald nimmt eine Zählung der Ich-Angaben vor); auch tritt oft eine ziemliche Fehlerhaftigkeit der Angaben auf. Der beschreibende Typ, welcher sich direkt der Sache zuwendet und wenig subjektive Selbständigkeit verrät, scheidet sich bei genauerer Prüfung in zwei ganz verschiedene Typen: den „passiven" und den „vorsichtigen". Der geringe Anteil

[1]) Obige Inhaltsangabe ist meiner „Differentiellen Psychologie" entnommen.

der Selbsttätigkeit beruht dort auf tatsächlicher Schwäche, hier auf einer bewußten Hemmung; der vorsichtige Typ ist daher durch viele Zweifel und Vorbehalte, sowie durch große Treue der Aussagen gekennzeichnet.

B. trifft also bei dem objektiven Typ bereits die wichtige Scheidung in die beiden Wertstufen, beim subjektiven fehlt noch die gleiche Teilung; doch ist das, was wir die höhere Form des subjektiven Typs nannten, wohl identisch mit dem „harmonischen" Typ B.s, in dem er eine Synthese des selbsttätigen und beschreibenden Typs sieht. Dieser harmonische Typ soll sich nach B. darin äußern, daß zuerst, unter absichtlicher Hemmung objektiver Beteiligung, eine sachliche Darstellung gegeben wird, daß aber dann in abgeschlossenen Deutungen und Kritiken die individuelle Stellungnahme des Ich zum Ausdruck kommt.

Von der Art der Bearbeitung, welche die Intelligenztätigkeit an ihrem (psychischen oder physischen) Gegenstand vornimmt, schreibt sich die Scheidung in analytische und synthetische Intelligenz her[1]). Freilich gibt es wohl keine Denkleistung, in der nicht beides, Analyse wie Synthese, enthalten wäre; schon die einfache Auffassung eines Gegenstandes ist Analyse, indem aus dem unbestimmten Chaos der Sinneseindrücke ein gewisser Komplex herausgelöst und zum gesonderten Zielpunkt der Aufmerksamkeit gemacht wird; sie ist Synthese, indem die zugleich wahrgenommenen Elemente zur Einheit des Dinges zusammengefaßt und das Erlebnis zu einem transsubjektiven, „gemeinten" Gegenstand in Beziehung gesetzt wird. Je schwieriger die I.-Aufgabe, je verwickelter der Denkprozeß, um so mehr sind an ihm analytische und synthetische Einzelakte in buntem Wechsel beteiligt. Dennoch gibt es auch hier vorwiegende Tendenzen, die bald durch das Ziel der Denkaufgabe, bald durch die Anlage des Individuums bestimmt sind. Eine Reihe von I.-Forderungen stellt den Menschen vor ein zunächst ungegliedertes Ganzes und gibt ihm auf, aus ihm ein Element (oder deren mehrere) herauszulösen: er soll aus einem gegebenen Inhalt den Begriff, den Sinn oder Widersinn, die Pointe,

[1]) Über die Bedeutung von Analyse und Synthese im Gesamtsystem des persönlichen Lebens vgl. Die menschliche Persönlichkeit, S. 139 f. Die folgende Stelle sei hier wiedergegeben: „Nur dadurch vermag die Person das Chaos der Welt zu bewältigen, daß sie aus ihrer Unendlichkeit Einheiten (Einzelobjekte, Einzelmerkmale, Einzelvorgänge) herauslöst und auf sie ihre Zielstrebigkeit konzentriert; und nur dadurch vermag sie wiederum mit dem gleichgültig und zufällig Nebeneinanderbestehenden fertig zu werden, daß sie die getrennten Einheiten zu gewissen Verbundenheiten vereinigt. Aber jenes Herauslösen ist nicht ein bloßes Zerstückeln und dieses Verknüpfen nicht ein bloßes Aneinanderreihen; sondern jede Analyse und Synthese und vor allem das Ineinander beider ist und bleibt von der Ganzheit aus zielstrebig bestimmt."

den Wert, die Fehlerhaftigkeit herausfinden; er soll das Objekt unter bestimmten Gesichtspunkten (der Form oder der Farbe, der Schönheit oder der Güte, der Handlung oder des Zustandes) betrachten; er soll zerlegen, beurteilen, kritisieren. Das sind in erster Linie analytische Akte. Andere I.-Aufgaben gehen dagegen aus von einer zunächst bloß aggregativen Mehrheit, deren Elemente nun in einen sinnvollen Zusammenhang gebracht werden sollen. Das Kombinieren, das zu getrennten Gegenständen einen Beziehungsgedanken ergänzend hinzufügt, das Vergleichen zweier oder mehrerer Elemente und die Gewinnung des übergeordneten Begriffs oder Gesetzes, das Herstellen einer Ordnung, das Erdenken eines Systems, das Entwerfen eines Planes, das Organisieren — alle diese Betätigungen sind vorwiegend synthetischer Natur.

Gewiß heißt es auch hier: jedes an seiner Stelle; wo kritische Beurteilung von Einzelheiten notwendig ist, soll man nicht mit weit ausgreifenden Zusammenfassungen kommen — und umgekehrt. Aber daß verschiedene Menschen in ganz verschiedenem Maße zu der einen oder anderen Art intellektueller Arbeit neigen, bleibt trotzdem bestehen. Es gibt analytische Naturen, deren Richtung und Stärke sich vor allem in der isolierenden Betrachtung von Einzelheiten, im Grenzenziehen und Fehleraufdecken (Kritisieren) bekundet; ihnen stehen die Synthetiker gegenüber, die darauf eingestellt sind, Getrenntes zusammenzuschauen und zusammenzudenken, Entferntes zu verbinden und Chaotisches zu ordnen, denen aber oft die Treue im Kleinen und die Schärfe im Einzelurteil fehlt [1]).

Diese Typik der I. hängt — worauf schon Meumann (*9*, *10*) aufmerksam gemacht hat — mit gewissen Dispositionen der Aufmerksamkeit eng zusammen. Man sieht hieran, daß die Neigungen zu Analyse und Synthese nicht nur Prinzipien der Denktätigkeit, sondern Grundformen persönlicher Lebenstätigkeit überhaupt sind.

---

[1]) Die analytische bzw. synthetische Geistesrichtung kann in verschiedener Zuspitzung ausgeprägt sein; zuweilen tritt sie so einseitig auf, daß sie in ihrer Begrenzung schon mehr als Sonderbegabung denn als I. anmutet. Indessen darf man doch nicht im eigentlichen Sinne von einem „Talent zur Analyse“ bzw. „Talent zur Synthese“ sprechen. Denn die hierin ausgedrückten Fähigkeiten erstrecken sich doch auf beliebige Inhaltsgebiete, sind also formale Fähigkeiten, während das „Talent“ durch die Beschränkung auf ein bestimmtes Inhaltsgebiet (Musik, Technik usw.) gekennzeichnet ist.

Die letzte Gliederung, die wir zu besprechen haben, steht am unmittelbarsten mit der Wirklichkeit des Alltags, insbesondere mit der beruflichen und sozialen Besonderheit der Individuen, im Zusammenhang: es ist die Unterscheidung von theoretischer und **praktischer Intelligenz.**

Wir Gelehrten sind nur allzu leicht geneigt, I. hauptsächlich in jenen Verhaltungsweisen zu suchen, die uns auf Grund unseres Berufes am nächsten liegen, also in der Bewältigung geistiger Aufgaben, deren eigentliches Ziel die Hervorbringung eines Denkergebnisses und deren Hauptmittel die sprachliche Formulierung ist. Die Beziehung zur handgreiflichen Wirklichkeit der Gegenstandswelt ist bei dieser „theoretischen" I. nur eine mittelbare; gewiß soll der weitausschauende Plan des Feldherrn, der Forschungsbefund des Gelehrten auf die Praxis des Lebens einwirken; aber der eigentliche I.-Akt liegt doch nicht in dieser Einwirkung, sondern in den geistigen Vorbereitungen hierzu.

Demgegenüber beginnt die I.-Forschung erst in letzter Zeit die selbständige Bedeutung der „praktischen" I. zu beachten, deren Betätigung unmittelbar am tätigen Leben angreift. Auch für sie gilt selbstverständlich unsere Definition der geistigen Anpassungsfähigkeit an neue Situationen; aber die Situation fordert hier einen direkten Eingriff in die gegebenen Gegenstandsbedingungen; ihre Neuheit verlangt, daß man „sich zu helfen weiß" (Lipmann) (*29*), und die Anpassung besteht in einer äußerlich sichtbaren Handlung; der ganze Vorgang spielt sich in der unmittelbaren Gegenwart und der sinnlichen Anschaulichkeit ab.

Die überwiegende Mehrheit der vom Menschen geforderten I.-Handlungen gehört diesem praktisch anschaulichen Gebiet an; und vermutlich wird auch die überwiegende Mehrheit der Menschen, die als „intelligent" gelten dürfen, viel mehr praktische als theoretische I. besitzen. In der Art, wie ein einsamer Wanderer im Augenblick einer Gefahr sich benimmt, oder wie ein Handwerker bei einem neuartigen Auftrag die zweckentsprechendsten Arbeitsmittel und Handgriffe wählt, oder wie ein junger Mensch, der zum ersten Male allein eine selbständige, verantwortungsvolle Handlung vollziehen muß, sich zu helfen sucht, oder wie eine Hausfrau beim Fehlen bestimmter Nährmittel durch Heranziehung von Ersatzstoffen und durch neue Kombinationen die Familie schmackhaft und nahrhaft zu versorgen weiß — in diesen praktischen Verhaltungsweisen kann nicht minder starke und wertvolle I. zutage

treten, als in Leistungen des Verstehens und Urteilens, des Definierens und Kritisierens, des sprachlichen Erfassens und Darstellens.

Die sprachlich-logische I. ist feinerer Schattierungen fähig; sie vermag scharfe Zuspitzungen zu vollziehen und leise Verschiedenheiten zu bewältigen — man denke etwa an die Fähigkeit, in einer Kultursprache durch die Fülle der Relationsausdrücke und Flexionsformen allen möglichen Schattierungen der logischen Beziehungen gerecht zu werden; sie vermag ferner mit dem Mittel der Abstraktion von der individuellen Zufälligkeit der Gegebenheit zum Allgemeinen vorzudringen, und mit Hilfe geistiger Rückschau und Vorausschau sich von dem bloßen Gegenwartsaugenblick zu befreien. Hierin liegen ihre Vorzüge, die nicht verkannt werden; sind sie es doch, die das spezifisch Menschliche an der Intelligenz ausmachen; denn Tiere sind wohl der praktisch-anschaulichen I.-Handlungen bis zu gewissen Grenzen fähig, niemals aber sprachlich-logischer I.-Betätigungen. Dennoch aber müssen wir den Intellektualismus abschütteln, der in solch theoretischem Tun allein wahre I.-Äußerungen erblickt und die in praktischer Tat sich bekundende I. übersieht. Diese ist in ihren Linien gröber und einfacher, sie ist enger gebunden an Anschauung und Gegenwart; aber innerhalb dieser Grenzen leistet sie Vollgültiges; und in tausend Fällen des Alltagslebens — in denen die theoretische I. versagt oder nur mühsam und auf Umwegen zum Ziel kommt — ist praktische I. das, was nottut, und was wir kennen, würdigen und pflegen müssen.

Noch kaum bearbeitet ist die Frage, inwieweit theoretische und praktische I. auf selbständigen, voneinander relativ unabhängigen Anlagen beruhen. Wir wissen zunächst nur aus den Erfahrungen des Alltags, daß die I. als Fertigkeit recht einseitig ausgebildet sein kann: es gibt hochintelligente Gelehrte, die in den Realitäten des Lebens kindlich hilflos sind; und es gibt hochintelligente Praktiker, die völlig versagen, wenn sie vor allgemeinen Denkaufgaben oder vor der Forderung exakter sprachlicher Formulierungen stehen. Im allgemeinen dürfte wohl Lipmann (*29*) im Recht sein, wenn er keine allzu hohe Korrelation zwischen theoretischer und praktischer I. annimmt, wohlgemerkt, solange I. hier als fertige Eigenschaft gefaßt wird; wie es mit der ursprünglichen Fähigkeit als solcher steht, bleibt noch ganz unentschieden. Denn man vergesse nicht, daß die Hinwendung zu Theorie oder Praxis in erster Linie eine Sache des Interesses und der Gemütsrichtung ist.

Es ist sehr wohl möglich, daß diese Interesseneinstellung im Menschen primäre Bedeutung hat und daß er dann seine (zunächst indifferente) I.-Anlage gemäß dem vorwiegenden Interesse übt und einseitig ausgestaltet. Größere Klarheit über diese Frage werden wir wohl erst aus Entwicklungsgeschichten kindlicher und jugendlicher Intelligenzen gewinnen, wie wir sie zurzeit noch nicht besitzen.

Merkwürdigerweise sind es nicht die Psychiater gewesen, die den so wichtigen Unterschied zwischen theoretischer und praktischer I. herausstellten; vielmehr hat die Psychologie sich aus eigener Arbeit — freilich auch spät genug — zu dieser Scheidung durchgearbeitet. Bestimmend wurden drei Forschungsgebiete, in denen sich die bloße Beachtung der „theoretischen" I. als unzulänglich erwies: die Tierintelligenz, die frühe Kindheit und die Eignung für praktische Berufe. Wie sehr unsere Untersuchungsmethodik unter dieser Einseitigkeit litt und zum Teil noch leidet, kommt im Kapitel VI zur Sprache. In neuester Zeit haben O. Lipmann (*29*) und E. Stern (*38*) mit besonderem Nachdruck auf den Unterschied der beiden I.-Arten hingewiesen. Wenn Lipmann aber hierbei die praktische I. als „natürliche" und „lebensnahe" bezeichnet, so scheint mir dies bedenklich. Theoretische I.-Handlungen sind für gewisse Lebensaufgaben ebenso angemessen und in diesem Sinne „natürlich" wie praktische I.-Handlungen für andere Aufgaben. Unnatürlich und lebensfern wirkt die theoretische I. nur in Situationen, wo sie nicht hingehört, und bei Menschen, deren Lebensform nicht auf sie eingestellt ist. Lediglich als methodische Anweisung für die künftige I.-Untersuchung darf man daher jene wertenden Bezeichnungen auf die beiden I.-Arten anwenden.

2. Intelligenzgrade.

Eine quantitative Bewertung der I. kann in doppelter Bedeutung erfolgen, je nachdem man die I. als Fähigkeit oder als Fertigkeit auffaßt. Im ersten Falle geht man von der Annahme aus, daß der Anlage als solcher eine Stärke innewohne, die als konstanter Antrieb hinter aller Entwicklung und Betätigung im einzelnen steht. Im zweiten Fall beachtet man die jeweils erreichte Höhe der I., die augenblickliche Fertigkeit im Vollziehen intelligenter Handlungen, einen Maßwert, der sich natürlich während der ganzen Zeit der Entwicklung im positiven Sinne ständig ver-

ändert. Dort wird also der geistige Habitus, hier der Status praesens mit einem Maßzeichen versehen. Jenen wollen wir den Intelligenzgrad, diesen die Entwicklungsstufe der Intelligenz nennen. Zunächst ist von dem I.-Grade zu handeln.

Die Frage, ob die I.-Anlage in verschiedenen Stärkegraden vorkomme, wird meist unbedenklich bejaht. Ja die oberflächliche Betrachtung ist geneigt, im Vergleich der Intelligenzen überhaupt keine anderen als gradweise Verschiedenheiten anzuerkennen, da sie die I. als einförmige, der Art nach nicht variierbare Eigenschaft ansieht.

Anders liegt die Sache für uns, die wir soeben feststellten, daß die I. in sehr verschiedenen Qualitäten vorkommen kann. Verhindert diese Buntheit der I.-Arten und -Typen denn nicht, daß man alle I. gleichmäßig abstufe und demnach jeder einzelnen I. einen Rangplatz in einer Stufenleiter der I.-Grade anweise? Mir scheint dies trotzdem möglich zu sein, aber nur dadurch, daß eine Reihe von Teilerscheinungen der I. einer gradmäßigen Wertung zugänglich sind, so daß aus ihnen ein Gesamtindex abgeleitet werden kann. Das Verfahren dieser Ableitung kommt erst an späterer Stelle zur Sprache (Kap. VII); hier aber müssen wir feststellen, in welchem Sinne man bei Teilerscheinungen der I. und infolgedessen bei der I.-Anlage selbst von Graden sprechen darf.

Jede Fähigkeit ist teleologisch bestimmt; der Grad einer Fähigkeit kann daher lediglich auf ihr Verhältnis zu den Zwecken, denen sie dient, bezogen werden. Dient die Fähigkeit verschiedenen Teilzwecken, so ist sie höher einzuschätzen, wenn der Teilzweck, auf den sie besonders eingestellt ist, innerhalb des Wertsystems der Zwecke höher steht. Innerhalb jedes Teilzwecks aber gibt es verschiedene Grade in der Annäherung an eine möglichst vollkommene Verwirklichung des Zweckes. So erhalten wir zwei Gesichtspunkte der Graduierung, den axiologischen und den dynamischen.

Die Hineinnahme von Wertungen in eine psychologische Untersuchung mag zunächst befremden. Und in der Tat, die rein phänomenologische Betrachtung hat kein Recht dazu. Sobald man aber nicht mehr bloß psychische Phänomene und Akte untersucht, sondern Fähigkeiten, sind wir auf Wertgesichtspunkte angewiesen. Denn Fähigkeiten sind die Mittel, durch welche die Persönlichkeit die ihr wertvollen Zwecke zu verwirklichen strebt; ohne Maßstäbe für diese Werte haben wir daher auch keine Maßstäbe für die Stärke der Fähigkeit.

Nun ist es an dieser Stelle ausgeschlossen, die gesamte Wertskala des persönlichen Lebens zu entwickeln und zu begründen. Wir können uns zum Glück mit der einfachen Annahme begnügen, daß der Wert eines Teilzwecks und der ihm dienenden Fähigkeit in Parallele steht zu der geistigen Entwicklungsstufe, auf der die Fähigkeit normalerweise auftritt. Die in der Entwicklung hintereinander erscheinenden und auseinander hervorgehenden Stadien der I. (vgl. den folgenden Abschnitt) wären danach zugleich — sofern sie bei verschiedenen Individuen nebeneinander auftreten — ebenso viele Gradstufen der I.

Überschauen wir unter diesem Gesichtspunkt die früher erörterten I.-Arten, so ergibt sich:

Die reaktive I. steht tiefer als die spontane, die praktische I. tiefer als die theoretische; denn das erste Glied jedes Vergleichungspaares findet sich bereits beim Tier, beim kleinen Kind, beim Naturmenschen; spontane und theoretische I.-Betätigung ist dagegen dem Menschen allein, und zwar in seiner ausgereiften Form vorbehalten.

Nicht ganz so eindeutig ist die Bewertung bei den anderen I.-Arten.

Analytische und synthetische I. können jede für sich zu hochwertigen Ausprägungen führen, die ihrer Bedeutung nach nicht gut aneinander gemessen werden können. Dagegen ist dies noch bei einfacheren Leistungen möglich. Zeigt doch die psychische Entwicklungsgeschichte, daß die Fähigkeit zu scharfer Isolierung und Erfassung von Einzelheiten schon auf Stadien eintritt, auf denen die Fähigkeit zur Verknüpfung und zur Herstellung von Zusammenhängen noch sehr wenig ausgebildet ist. So sind die I.-Akte des Tieres und des kleinen Kindes zweifellos viel stärker analytisch als synthetisch gefärbt. Auch setzt Synthese in viel höherem Maße vorgängige Analyse voraus als umgekehrt. Somit kann für einfache und mittlere I.-Stufen die synthetische Leistungsfähigkeit im allgemeinen als der analytischen Leistungsfähigkeit überlegen gelten.

Der Unterschied zwischen objektiver und subjektiver I. läßt sich überhaupt nicht auf ein Mehr oder Minder bringen; vielmehr fanden wir, daß sich innerhalb jeder von beiden Arten eine minderwertige und eine höherwertige Form feststellen läßt. Unbestreitbar am höchsten steht die selten vorkommende harmonische Verbindung des objektiven und subjektiven Verhaltens, der „harmonische Typ" Baerwalds (s. S. 21).

Der dynamische Gesichtspunkt der Gradabstufung bereitet weniger Schwierigkeiten. Hier handelt es sich um Vergleiche der Energieentfaltung innerhalb gleichartiger Betätigungsweisen, und da können Maßstäbe zugrunde gelegt werden, die selber der quantitativen Abstufung unterliegen. So ist unter sonst gleichen Bedingungen (der Aufgabe, des Alters und der Umwelt) diejenige I. die höhere, die bestimmte Leistungen mit größerer Schnelligkeit, oder mit geringerer Fehlerhaftigkeit, oder in weiterem Umfange, oder in größerer Vielseitigkeit vollbringt. Es ist ferner diejenige I. höher einzuschätzen, die aus einer nach der Schwierigkeit geordneten Aufgabenreihe die schwerere zu lösen vermag usw.

Hierbei ist zu beachten, daß jeder dieser Vergleichsgesichtspunkte nur „ceteris paribus" gilt. Größere Schnelligkeit des intelligenten Handelns ist an sich noch kein Zeichen höherer I. — wenn sie nämlich erkauft ist durch größere Flüchtigkeit, Fehlerhaftigkeit, Unzuverlässigkeit. Nur dann, wenn gleich gute Leistungen von dem einen bedeutend schneller als von dem anderen vollzogen werden, besteht ein Recht, bezüglich dieser Teilseite der I. von einem höheren Grade zu sprechen.

Der Gesamtgrad der I. eines Menschen ergibt sich nun als Resultante aus den Teilabstufungen, die wir eben besprachen. Aber es ist dies nicht eine bloß mosaikartige Zusammensetzung; vielmehr vereinigen sich die Teilfähigkeiten zu einer einheitlichen Beschaffenheit, mit der sich der Mensch in der Welt geistig zurechtfindet, zu einer Gesamtwertigkeit seines geistigen Anpassungsvermögens an die Neuforderungen des Lebens. In dieser Gesamtwertigkeit haben nicht alle Teilmomente gleiche Größe; vielmehr findet hier eine weitgehende Kompensation statt; eine geringere Fähigkeit zu analytischem Tun wird durch größere Fähigkeit zur Synthese wett gemacht; was der eine durch Schnelligkeit erreicht, gelingt dem anderen durch Gründlichkeit usw. So wechseln Wellenberge und -täler in mannigfacher Weise; dennoch hat das Niveau, auf dem sich im ganzen die geistige Auseinandersetzung und Abfindung mit der Welt bewegt, eine Gesamthöhe oder -tiefe, die in eine Gradziffer gefaßt werden kann.

Man darf nun freilich niemals glauben, durch eine solche Gradangabe die I. eines Menschen erschöpfend charakterisiert zu haben; daneben behält die qualitative Charakteristik ihren vollen selbständigen Wert.

IV. Kapitel.

Die Entwicklung der Intelligenz.

Die Abhängigkeit der I. vom Alter (s. S. 14) bedingt, daß die Anlage sich allmählich in einer Abfolge von Leistungsmöglichkeiten entfaltet, die einer inneren Gesetzmäßigkeit gehorcht. Gerade für uns, die wir die kindliche und jugendliche I. erforschen sollen, ist diese Berücksichtigung der Entwicklungstatsachen unumgänglich. Hierbei begegnen uns wieder die beiden Gesichtspunkte der Quantität und Qualität, nur daß innerhalb jedes dieser Gesichtspunkte jetzt nicht mehr, wie im vorigen Kapitel, das Nebeneinander bei verschiedenen Individuen, sondern das Nacheinander innerhalb identischer Individuen zur Erörterung steht, also die quantitative bzw. qualitative Veränderung der I.

Jede psychische Entwicklung ist Wachstum und Wandlung zugleich; und so betätigt sich auch die I. mit steigendem Alter nicht nur in immer stärkeren Kraftäußerungen, sondern auch in qualitativ wechselnden Formen. Manche dieser chronologischen Tatsachen sind ja zweifellos außenbedingt (durch Unterricht, Vorbild usw.); aber es bleiben doch genug Züge übrig, die in der allmählichen Verwirklichung der Anlage selbst ihren Grund haben.

1. Wachstum der Intelligenz.

Beginnen wir diesmal mit den quantitativen Erscheinungen, so zeigt sich, daß jede dynamische Seite der I. im Laufe der Jahre eine Steigerung erfährt. Es wächst die Geschwindigkeit, mit der bestimmte geistige Akte vollzogen werden können; es wächst bei gleichbleibender Schwierigkeit der Aufgabe die Leichtigkeit in der Bewältigung und die Güte der Leistung; und wenn eine Reihe verschieden schwerer Aufgaben gestellt wird, so vermag die I. mit steigendem Alter immer größerer Schwierigkeiten Herr zu werden.

Dieser Altersfortschritt bei bestimmten Aufgaben ist freilich nur so lange wirksam, als die Anforderungen in der Tat eine merkliche „Schwierigkeit" haben; man darf deshalb nicht erwarten, mit Aufgaben, die für jüngere Jahrgänge geeignet sind, bei höheren noch wesentliche Entwicklungsunterschiede nachzuweisen. Eine Nichtbeachtung dieses Gesichtspunkts hat oft zu methodischen Mißgriffen geführt. Wo die Anforderung zu leicht wird, tritt die eigentliche I. überhaupt nicht mehr in Funktion; und man prüft ganz

anderes, als man zu prüfen glaubt. Entwicklungspsychologisch sind gerade solche Feststellungen sehr lehrreich; indem sie uns nämlich zeigen, in welchen Altersstufen bestimmte Leistungen einen besonders starken Altersfortschritt zeigen und in welchen sie dann zu einem relativen Stillstand kommen, wird jeder Leistungsform ihre eigentliche Reifungszeit innerhalb der geistigen Gesamtentwicklung zugewiesen. Besitzen wir erst eine größere Anzahl solcher Kurven des Altersfortschritts für verschiedene I.-Funktionen, dann werden wir auch ganz anders als heute ein wirkliches Bild der typischen I.-Entwicklung durch die ersten zwei Lebensjahrzehnte hindurch zeichnen können; es wird sich dann ergeben, daß das Wachstum der I. nicht in einem gleichmäßigen Anstieg aller Funktionen, sondern in der Aufeinanderfolge verschiedener jeweilig im Vordergrund stehender Leistungsprozesse beruht. Zugleich werden wir für jede Altersstufe die ihr gemäßen Anforderungen einwandfrei festlegen und daraus Richtlinien für psychologische Prüfungen und pädagogische Einwirkungen ziehen können.

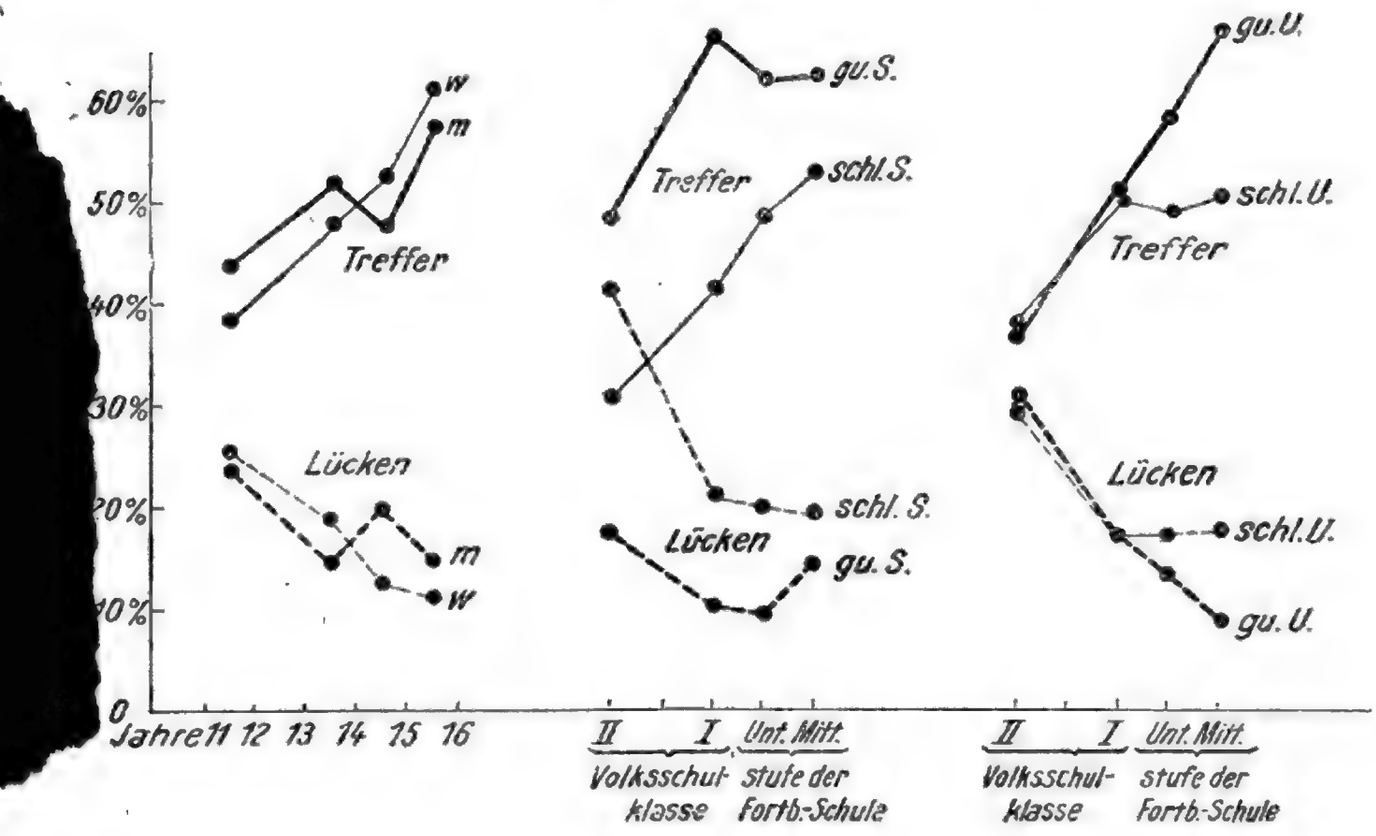
Von Untersuchungen, die über solche bloß methodische Absichten hinaus in die Struktur und die Bedingungen des Altersfortschritts selbst Einsicht verschaffen sollten, seien zwei aus dem Breslauer Institut stammende erwähnt.

Bei Aussageversuchen mit einem Bilde wurden von mir Kinder und Jugendliche von sieben bis zu achtzehn Jahren geprüft (*40*, S. 101 u. 120). Untersucht wurde die Fähigkeit: 1. selbständig den Inhalt des gesehenen Bildes anzugeben („Bericht"), 2. auf Befragen noch weitere Angaben über behaltene Einzelheiten zu machen („Verhör"). Es stieg nun die Menge der im Bericht erzielten Angaben von 7 bis zu 14 Jahren um mehr als 100 %, dann bis zu 18 Jahren noch um weitere 24 %, während die Anzahl der Verhörsangaben von 7 bis zu 14 Jahren nur um 21 %, später gar nicht mehr stieg. Nun verlangt die selbständige Wiedergabe des Gesehenen viel mehr spontane Denkarbeit, während die Beantwortung von Verhörsfragen eine viel mehr rezeptive Leistung ist. Es ergibt sich also, daß die Spontaneität einen viel stärkeren und viel längeren Altersfortschritt zeigt als die Rezeptivität. In gleicher Richtung liegt das Ergebnis, das innerhalb des Verhörs die Antworten auf Suggestivfragen einen weit stärkeren Fortschritt aufwiesen als die Beantwortung der gewöhnlichen Fragen: der Widerstand gegen Suggestion ist also ebenfalls ein Faktor, der aufs deutlichste mit dem Alter wächst, und zwar in besonders hohem Maße bis zum 14. Jahre.

Andere Seiten der Intelligenz prüfte Minkus (*31*) in seiner Massenuntersuchung an sieben Jahrgängen von Volks- und Fortbildungsschulen.

Infolge des frühen Todes von Minkus konnte bisher nur ein kleiner Bruchteil der Ergebnisse (die vier Klassenstufen von 12 bis zu 15 Jahren in bezug auf den einen Test: Ergänzung fehlender Bindewörter in einem ziemlich schwierigen Text) bearbeitet werden. Die von mir durchgeführten Berechnungen ergaben, in verschiedenen Gruppenteilungen, die in Fig. 1, 2, 3 dargestellten Bilder des Altersfortschrittes. Gezählt wurde, wieviel Prozent der im Text enthaltenen Lücken

richtig ausgefüllt wurden („Treffer") und wieviele Lücken unausgefüllt blieben („Lücken"). — Die erste Figur zeigt bei beiden Geschlechtern einen starken Altersfortschritt, der nur bei den Knaben (m) in der Mitte (beim Übergang von der Volksschule zur Fortbildungsschule) eine Unterbrechung zeigt. — Fig. 2 trennt die guten von den schlechten Schülern mit dem bemerkenswerten Erfolg, daß der Altersfortschritt bei den guten Schülern mit dem 14. Jahr aufhört, bei den schlechten bis zum 16. Jahr weitergeht. Für die Begabten wird also der Test schon mit 14 Jahren zu leicht, als daß er noch ein typisches Symptom der I. lieferte, für die schwachen ist seine Schwierigkeit noch groß genug, um eine allmählich steigende Bewältigung herbeizuführen. — Fig. 3, die nicht nach Fähigkeiten, sondern nach den sozialen Umweltverhältnissen gliedert, bietet sonderbarerweise das entgegengesetzte Bild: hier sind es die aus guten sozialen Ver-



hältnissen stammenden älteren Schüler, die noch Leistungszunahmen zeigen, wo die sozial schlechter gestellten nicht mehr fortschreiten. Dies ist wohl so zu deuten: je leichter eine Aufgabe wird, um so weniger ist ihre Lösung von der Begabung, um so mehr von den unterstützenden Bedingungen abhängig, die die Umwelt (durch gehobene Sprachkultur usw.) bietet. (Zum letzten Punkt vgl. Kap. X.)

Unter den dynamischen Bedingungen der I., die mit steigendem Alter zunehmen, ist vor allem die Stärke der determinierenden Tendenz zu nennen. Bei jeder Intelligenzleistung ist der Ablauf des psychischen Geschehens durch eine — selbstgesetzte oder von außen erhaltene — Aufgabe bestimmt; der Zielgedanke — die Absicht, der Plan — muß also eine herrschende Rolle gegenüber der Fülle der Bewußtseinsphänomene spielen, muß die Aus-

lese und die Richtung ihrer Verknüpfung „determinieren". Sobald nun die I.-Leistung nicht momentan erledigt werden kann, ist ihre Lösung davon abhängig, daß die determinierende Tendenz die nötige Breite, Gliederung und Ausdauer hat. Breite: Sie muß eine Mannigfaltigkeit von Bewußtseinsinhalten und -abläufen in Betracht ziehen, um die dem Ziel angemessene Auslese zu treffen. Gliederung: Ein verwickelterer I.-Akt kann sein Endziel nur schrittweise erreichen; einzelne Mittel, die dem Gesamtzweck dienen, werden zu Zwischenzielen, die durch determinierende Tendenzen niederer Ordnung erreicht werden (es müssen Hindernisse beseitigt, Hilfswerkzeuge beschafft, Teilaufgaben gelöst werden); aber bei alledem darf die übergeordnete Zielrichtung des Gesamtprozesses nicht aus den Augen gelassen werden. Ausdauer: Höhere Intelligenzakte erstrecken sich über längere Zeiten, erfordern ein längeres Beharren bei gewissen Einstellungen, Widerstandsfähigkeit gegen Ablenkungen, die Fähigkeit des Wiederaufnehmens von Fäden, die vorübergehend fallen gelassen worden waren.

Hier stehen wir wieder an einer Stelle, wo die I. mit anderen psychischen Funktionen aufs engste zusammenhängt. Denn die Kraftverhältnisse der determinierenden Tendenz sind nicht alleinige Eigentümlichkeiten der geistigen Leistungsfähigkeit, ja sie sind primär überhaupt nicht geistiger Art, sondern Merkmale der Willens- und Aufmerksamkeitssphäre; und die beste geistige Anlage bleibt ein Schattengebilde, wenn ihr nicht durch die in der determinierenden Tendenz hervortretenden Willensbeherrschung Lebensblut zugeführt wird.

Die Steigerung dieser Determinationsfähigkeit ist aber eine der wichtigsten Entwicklungstatsachen; sie ermöglicht es, daß das Kind von einfachen, im Augenblick zu erledigenden I.-Aufgaben allmählich zu solchen Aufgaben aufsteigt, die verwickelter sind, planmäßige Gliederung und längere Zeit in Anspruch nehmen. Das Versagen jüngerer Altersstufen bei I.-Aufgaben beruht oft genug nicht etwa darauf, daß die notwendigen geistigen Vorbedingungen des Verstehens, Kombinierens usw. nicht verfügbar wären, sondern darauf, daß die determinierende Tendenz nicht durchhält und die zunächst richtig einsetzende Verwertung jener ständigen Hilfsmittel nicht zu Ende gebracht wird. Willenserschlaffung und vorzeitiges Nachlassen der Aufmerksamkeitsspannung bewirken dann, daß der intelligenten Einleitung einer Handlung nicht ihre intelligente Durchführung entspricht.

2. Wandlungen der Intelligenz.

Die qualitativen Wandlungen, welche die Intelligenz im Laufe der Entwicklung durchmacht, kann man sich am besten vergegenwärtigen, wenn man die primitivsten Formen ihres Auftretens mit reiferen Formen vergleicht. Hier haben die neuesten Forschungen auf dem Gebiet der frühkindlichen Psychologie einerseits, der Tierpsychologie andererseits unseren Blick bedeutend geschärft (während auf dem dritten Gebiet seelischer Frühstadien, der I. der primitiven Völker, die psychologische Forschung noch in bedauerlichem Rückstand ist)[1]). — Unsere folgende Betrachtung verzichtet absichtlich darauf, die Phänomenologie der primitivsten Denkprozesse bei Mensch und Tier zu erörtern; dafür muß auf die besondere Literatur verwiesen werden; uns kommt es lediglich auf die teleologische Beschaffenheit der frühen Intelligenz an.

Wir wissen heute, daß auf allen genannten Frühstufen echte intellektuelle Prozesse vorkommen; schon in den ersten Jahren der Kindheit und ebenso im Leben der höheren Tiere besteht die Fähigkeit zu Leistungen, die nicht durch bloße Vorstellungsbildung, -nachwirkung und -verknüpfung zu verstehen sind, die nicht auf Instinkt, Gewöhnung, Übung und Gedächtnis zurückgeführt werden können, sondern die eine auf neue Lebensbedingungan eingestellte Neutat bedeuten. Aber diese Frühintelligenz ist noch fast oder ganz ausschließlich reaktiver Natur, abhängig von den äußeren Reizen in ihrer Augenblickskonstellation; es fehlt Vorschau und Vorwegnahme von künftigen Antrieben und Entscheidungen von innen heraus, Berücksichtigung von nicht sinnlich Gegebenem. Es fehlt ferner die eigentliche Intellektualität, d. h. Freude am Geistigen um seiner selbst willen; die ersten I.-Leistungen sind nichts anderes als Umwege der Instinktbefriedigung, die dann gesucht werden, wenn die geradlinige Erreichung instinktiver Ziele auf Hemmungen stößt, insbesondere wenn die Nahrung dem hungernden Individuum nicht auf normalem Wege zugänglich ist. Die Frühintelligenz bedarf des Antriebs durch starke elementare Gefühle, die durch Handlungen befriedigt werden sollen; sie ist durch und durch praktische I.

---

[1]) Das frühkindliche Denken ist ausführlich behandelt in folgenden Büchern: Groos (*7*); Stern (*13*); Bühler (*6*). Bezüglich der tierischen I. sind von entscheidender Bedeutung die Untersuchungen von Köhler (*26*).

Noch vor Vollendung des ersten Lebensjahres sind Säuglinge imstande, die von ihnen gehaltene Milchflasche in ihrer Stellung selbst zu regulieren, so daß sie um so steiler gehalten wird, je weniger Milch noch in ihr enthalten ist. Hier liegt fraglos eine Anpassung an die individuelle Beschaffenheit der jeweiligen Situation vor. — Einen ganz selbständigen, wenn auch mit untauglichen Mitteln ausgeführten Versuch zu einer Handlung, die sich völlig von den Geleisen des eingeübten Tuns entfernte, beobachteten wir bei unserem genau einjährigen Sohn: Als er die Flasche fast ausgetrunken hatte, so daß nur noch eine weiße Neige drinnen war, drehte er sie um und versuchte ihren Boden in den Mund zu stecken, als ob er glaubte, von dort aus den Rest erreichen zu können.

Vorzügliche Beispiele tierischer I.-Leistungen schildert Köhler (*26*). Er brachte vor den Augen seiner Schimpansen Bananen an Orten an, zu denen sie nicht auf dem üblichen Wege (durch Hinlangen mit den Armen oder durch Klettern) hinkommen konnten. Die Tiere wußten sich nun in ausgesprochen intelligenter Weise zu helfen: sie ergriffen einen Stock, um die Bananen durch die Gitter des Käfigs hindurch zu angeln; ja als sich ein Stock als zu kurz erwies, gelang es dem einen Affen, zwei Bambusstäbe ineinander zu schieben und damit sein Ziel zu erreichen. Wenn die Bananen am Dach des Käfigs befestigt waren, holten die Tiere eine Kiste herbei, schoben sie unter die Banane und kletterten hinauf usw. Die eingehende psychologische Analyse zeigte, daß es sich nicht etwa um blinde Zufallstreffer handelte, sondern daß die neuartigen Handlungen durch einen Zielgedanken determiniert, also „einsichtig" gewesen sind.

Aber die gleichen Versuche beweisen auch, wie eng das etwaige Hervorspringen des I.-Funkens beim Tier an die natürlichen Bedürfnisse und biologischen Daseinsbedingungen geknüpft ist. Unter diesem Gesichtspunkt erscheinen die behaupteten I.-Leistungen der Elberfelder Pferde (*27*) doppelt unwahrscheinlich; denn Rechen- und Buchstabierleistungen haben mit den natürlichen Lebensinteressen dieser Tiere nichts zu tun; sie stehen als isolierte Fremdkörper in dem biologischen Zwecksystem der Pferde. Der ganze diesen Tierexperimenten zugrunde liegende Gedankengang scheint schon im Ansatz falsch zu sein. Als ob I. so unabhängig wäre von dem gesamten Lebens- und Interessenkreis des Individuums, daß einfach die Ziele, die beim Menschen durch I. erreicht werden sollen (wie Rechnen und Buchstabieren), für das Pferd ebenfalls Ziele sein müßten! Gänzlich lebensfremde Zielsetzungen sind jedem Wesen, es sei Mensch oder Tier, nur durch Dressur anzuimpfen; und etwas Dressurähnliches wird deshalb auch bei den Elberfelder Pferden zugrunde liegen.

Die praktische I. des Kindes, die sich aus den eben geschilderten Anfängen entwickelt, bleibt noch lange Zeit ganz im Konkreten und Gegenwärtigen stecken; weitausschauende Pläne und Einschaltungen abstrakter Mittelglieder — etwa um bestimmte Hilfswerkzeuge und Hilfswege auf Grund allgemeiner Erwägung und Theorie zu finden — fehlen noch lange. Dagegen macht sich das Wachsen des Erfahrungskreises geltend, indem Wissensinhalte immer vielseitiger für neue Einstellungen verwertet werden können; und die Zunahme der determinierenden Tendenz macht immer kom-

pliziertere Handlungen möglich, in welchen schon eine Über- und Unterordnung von Hauptziel und Teilzielen besteht.

Von einem $2^3/_4$ Jahre alten Kind berichten wir an anderer Stelle (*13*, S. 296): Die Kleine sollte eine Fußbank an ihren Platz stellen. Dabei mußte sie einen Engpaß zwischen dem Sofa und einem nah gerückten Stuhl passieren. Sie sah, daß sie mit der Fußbank zusammen nicht durchkommen könne; da stellte sie die Fußbank auf den Stuhl, zwängte sich durch und hob nun die Fußbank vom Stuhl ab, um sie unter den nahen Tisch zu stellen.

Der qualitative Hauptfortschritt aber besteht bei der praktischen I. darin, daß sie nicht mehr bloß als Hilfsmittel elementarer Instinktbefriedigung auftritt. Sie gewinnt allmählich immer feinere Schwellen und bedarf nicht mehr so starker Anreize, um in Wirksamkeit zu treten. Es bilden sich zwei Gebiete dieser Wirksamkeit aus. Einmal das Gebiet realer Lebensanforderungen, die aber nun sehr viel mannigfaltiger sind als früher: bald muß das Kind sich in der Verlegenheit einer Situation, in der es allein gelassen ist, selbst zu helfen wissen; bald sucht es sich einer Strafe zu entziehen, indem es das Corpus delicti versteckt oder eine Ausrede ersinnt. Dann ist wieder eine Leckerei, eine schöne Spielsache das Ziel, dem seine intelligente Handlung zustrebt. In diesen Beispielen sind es noch immer starke Gefühle, die den Akt hervorrufen; daneben aber tritt in immer steigendem Maße eine kühlere, gefühlsärmere Art der praktischen I. auf, die gerade wegen ihrer geringeren Subjektivität um so größeren Praxiswert besitzt. Es ist das Gebiet des Auftrags: die Handlung ist nicht mehr durch ein gefühlsstarkes Motiv von innen her veranlaßt, sondern durch eine autoritative Aufgabestellung von außen her. Da soll ein Gegenstand aus einem anderen Zimmer geholt, in der Küche eine Bestellung ausgerichtet werden; das größere Kind muß Einkäufe machen, die jüngeren Geschwister betreuen, den Eltern im Geschäft helfen usw. Natürlich kann Übung und Gedächtnis hierbei bedeutend mitwirken; aber meist hat doch ein Auftrag seine individuelle Besonderheit; und die Art, wie sich das Kind eben mit diesen nicht übungsmäßig festgelegten Seiten seines Auftrags abfindet, das Maß dessen, was man ihm an Selbständigkeit der Entschließung, der Wahl usw. zumuten darf, bekundet seine praktische I.

Das andere Betätigungsgebiet der praktischen I. ist das Spiel. Wenn irgendwo, so bewährt sich der Vorübungswert des Spieles hier. Denn während die realen Lebensforderungen, mit denen das

Kind zu tun hat, sehr begrenzt sind und ernsthafte Aufträge, schon wegen des damit verbundenen Risikos, recht selten an es herantreten, bietet das Spiel eine unerschöpfliche Vielheit der Situationen und Aufgaben; denn jede Seite der Praxis mit den dazugehörigen Anforderungen: Gemeinschaftsleben und Kampf, Haus und Öffentlichkeit, die Tätigkeiten der Berufe und die Vorgänge der Zeitgeschichte — sie finden ihren Niederschlag im kindlichen und jugendlichen Spiel. Hierbei wird eine solche Vielseitigkeit praktischer I.-Leistungen entwickelt, daß dagegen die wenigen intelligenten Ernsthandlungen, zu denen das Kind Gelegenheit findet, recht dürftig erscheinen. Auch ist diese praktische Spielintelligenz viel spontaner, nicht erst durch Aufträge von außen hervorgelockt. Dagegen mangelt ihr ein Zug, der der ausgebildeten praktischen I. grundwesentlich ist: das Verantwortlichkeits- und Nützlichkeitsbewußtsein. Erst diejenige Handlung ist ja eine Bekundung echter geistiger Anpassung, die auch in ihren Konsequenzen der objektiven Tatsächlichkeit gewachsen ist; und jede intelligente Ernsthandlung reiht sich sinnvoll ein in die Verkettung der objektiven Weltzusammenhänge, weist über sich selbst hinaus auf einen angestrebten dauernden Nutzen. Die intelligente Spielhandlung weiß hiervon nichts, sondern beschränkt sich auf die isoliert bleibende Anpassung an die wirklichkeitsfremde Spielaufgabe, ist dafür aber auch frei von Utilitarismus. Hierin liegt vielleicht der qualitative Grundunterschied zwischen der praktischen I. des Kindes und der des Erwachsenen, und das drückt sich auch unwillkürlich in unserer Bewertung aus: ein Kind kann hochintelligent sein und doch seine Kindlichkeit bewahren; unkindlich wirkt es erst dann, wenn diese Intelligenz zu realistisch verwertet wird und damit vorzeitig aus der Schein- und Vorübungssphäre des Spiels in das kalte Licht des Ernstlebens gezerrt wird. Sehr fein hat die deutsche Sprache hierfür das Wort „Altklugheit" geprägt. Eine solche Verfrühung der I.-Entfaltung — die meist zugleich die Verkümmerung eines schwer zu entbehrenden Entwicklungsstadiums bedeutet — kann bald durch äußere Verhältnisse herbeigeführt sein (Heranziehung des Kindes zu den Unterhaltungen und Beschäftigungen der Erwachsenen, Kinderarbeit, Anhalten zum Bettel usw.), bald auf einer besonderen Anlage des Kindes beruhen, die leider noch von vielen Eltern als ein besonderer Vorzug aufgefaßt wird.

Zwischen dem reinen Spiel und dem vollen Lebensernst mitten-

inne steht das Gebiet der Handfertigkeit (das Zeichnen, Formen, Basteln, Konstruieren, Tischlern, Schneidern usw.); und gerade diese Zwischenstellung verleiht ihm seine besondere Bedeutung in der reiferen Kindheit, also in den Zeiten, die der Pubertät vorangehen und sie einleiten. Mit dem Spiel gemeinsam hat es die vorwiegende Freude an der Tätigkeit selbst, das Fehlen des vollen Verantwortungsbewußtseins für die mit dem Ergebnis verbundenen Konsequenzen; mit der praktischen Ernsthandlung gemeinsam hat es die Herstellung von Werken, die ein den Erzeugungsakt überdauerndes und vom Erzeuger losgelöstes Dasein führen und unter Umständen praktisch nutzbar gemacht werden können. Natürlich kann der Anteil des Spielmäßigen und des Ernsthaften je nach Umständen und nach dem Alter sehr verschieden sein. Soviel aber ist sicher, daß eine solche Mischung einen besonders lebhaften Anreiz zu intelligenten Handlungen bietet. Mit Recht wird daher in neuester Zeit darauf hingewiesen, daß Handfertigkeitsaufgaben ein besonders jugendgemäßes Mittel zur Schulung der I. seien, und daß das kindliche Verhalten zu solchen Aufgaben ein früher zu wenig beachtetes Symptom der I. darstelle. In der Tat beweisen viele Kinder, denen auf Grund ihrer Herkunft oder ihrer Veranlagung sprachlich-logische I.-Aufgaben weniger liegen, bei technisch-handwerklichen Aufgaben, daß sie einer wirklichen Neueinstellung auf neue Anforderungen, also eigentlicher I.-Handlungen fähig sind.

Die steigende Verwicklung der praktischen I.-Aufgaben bringt es mit sich, daß sie in zunehmendem Maße mit sprachlich-logischen Elementen durchsetzt werden; aber diese andere Seite der I. erfährt ihre besondere Entwicklung und gelangt schließlich dazu, sich als „theoretische" I. von der unmittelbaren Einwirkung auf die praktische Lebensgestaltung loszulösen und in dem Bewältigen rein geistiger Anforderungen aufzugehen.

Ihre erste Entwicklungsphase macht die sprachlich-logische I. beim Spracherwerb durch. Die Kindersprachforschung[1]) hat mit voller Sicherheit erwiesen, daß die Aneignung der Muttersprache kein bloßes „Lernen" im engeren Sinne des Wortes, also ein rein nachahmungsmäßiges Übernehmen und gedächtnismäßiges Behalten des von der Umgebung dargebotenen Sprachstoffes ist; vielmehr erarbeitet das Kind sich selbst seine Sprache an der Hand jenes

[1]) Gesamte Literatur bei C. und W. Stern (*43*).

Sprachstoffes nach Klang und Bedeutung; und diese Arbeit ist ein fortwährendes Neuentdecken von Beziehungen zwischen Laut und Sinn, ein Neugestalten von Gedanken dadurch, daß sie an der sprachlichen Formulierung ihr festes Knochengerüst erhalten, ein allmähliches Anpassen der zunächst noch ganz launenhaften Bildungen der Individualsprache an die festen Formen und Bedeutungen der Umgebungssprache — also I.-Leistung auf Schritt und Tritt.

Auch das Tier verwendet schon gewisse Laute als Ausdrucksbewegung für primitive Bewußtseinsinhalte; das Menschenkind aber schreitet über diese tierische Sprachgrenze in dem Augenblick mit Entschiedenheit hinaus, da es die gewaltige Entdeckung macht: daß jedes Ding einen Namen habe. Mit $1\frac{1}{2}$ Jahren beginnt es, auf die verschiedensten Gegenstände zeigend, zu fragen: is'n das? (was ist denn das?) und bekundet damit, daß ihm eine Ahnung von diesen drei Einsichten aufdämmert: Worte sind Symbole für Gegenstände; habe ich die Worte, so ist damit meine denkende Herrschaft über die Gegenstände errungen; zu jedem Gegenstand gehört ein solches Denkzeichen, und ich muß es mir zu erwerben suchen. So erarbeitet sich das Kind zunächst die gegenständliche Welt, dann die der Tätigkeiten (zu den Hauptwörtern, aus denen der primitive Wortschatz fast ausschließlich besteht, gesellen sich die Verben), endlich die Eigenschaften und Beziehungen. Im dritten Jahre werden die starren Einzelworte biegsam; die feineren Formabwandlungen der Begriffe: Einheit und Mehrheit, Zukunft und Vergangenheit, Steigerung und Minderung, werden durch Flexionsformen erfaßt. Und um weniges später beginnt das Kind auch, die Über- und Unterordnung der Satzgedanken sprachlich und damit auch logisch zu bewältigen: Satzgefüge mit Haupt- und Nebensätzen treten auf. Dieser Entwicklung parallel geht eine solche der fortschreitenden Sonderung des Objektiven und Subjektiven: die Wortbedeutungen, die zunächst Gefühls- und Willensmäßiges unklar und oft wechselnd mit sachlicher Konstatierung verbanden, werden zum weitaus größeren Teil objektiviert zu Bezeichnungen von Tatbeständen der Außenwelt, die vom Ich unabhängig sind; andere Worte werden daneben zu ausdrücklichen Zustands- und Gefühlsbezeichnungen des eigenen Ich geprägt. Dabei ist doch ganz wesentlich der geistige Blick nach außen gewendet; so ursprünglich und stark das Gefühls- und Willensleben im Kinde ist, zum Gegenstande der Selbstschau und zum Problem seines Denkens wird es

ihm erst sehr viel später. Darum geht die Entwicklung auch, nachdem die eigentliche Spracherlernung erledigt ist, vorerst im gleichen Sinne weiter. Noch gibt es genug da draußen zu fragen, zu erfassen, zu bewältigen; noch ist das Anschauliche durchaus der Interessenmittelpunkt. Das Kind forscht nach dem Was, Woher und Wozu, nach dem Wann und Wo. Es beginnt die Anschauungen der Außenwelt zu vergleichen und zu unterscheiden; es erkennt Ähnlichkeiten und reiht Gleichartiges aneinander — woraus sich dann nach mancherlei Mühen die Fähigkeit der Verallgemeinerung entwickelt. Die nun in ihren Hauptlinien beherrschte Sprache wird zum Mittel des Verstehens der von außen dargebotenen Gedankenzusammenhänge und zum Mittel der Darstellung selbsterzeugter Gedanken. Das Kind lernt an komplexen Erlebnissen gewisse Bestandteile oder Merkmale herausheben, und dann wiederum Vereinzeltes und Ungeordnetes zum Ganzen eines sinnvollen Zusammenhangs verknüpfen — von derartigen analytischen und synthetischen Leistungen der I. war ja schon oben (S. 21) die Rede.

In dieser geistigen Bewältigung der äußeren Wirklichkeit läßt sich nun eine kategoriale Gesetzmäßigkeit der Entwicklung feststellen; d. h.: die einzelnen Kategorien der Wirklichkeit werden in einer bestimmten Reihenfolge der denkenden Erfassung, Bearbeitung und Darstellung zugänglich. Und zwar beginnt die geistige Eroberung der Welt bei den Gegenständen; dann werden neben diesen die Handlungen und Vorgänge zum besonderen Ziel der Beachtung; endlich gehen auch die Eigenschaften und Zustände, sowie das vielgestaltige Gebiet der Beziehungen in das Denken ein. Ich habe seinerzeit (*40* und *42*) diese Stadienreihe mit den Worten „Substanzstadium", „Aktionsstadium", „Relations- und Merkmalsstadium" bezeichnet.

Dies Stadiengesetz soll nun nicht bedeuten, daß der Mensch als Ganzer sich zu einer gewissen Zeit in dem ersten Stadium befindet, dann zum zweiten übergeht usw.; vielmehr gilt die Abfolge für jede geistige Leistungsform, die es überhaupt mit der Erfassung und Bearbeitung der Außenwelt zu tun hat, besonders. Bei der Sprachentwicklung konnten wir die drei Epochen schon sehr früh feststellen (s. S. 38), bei anderen Leistungen, z. B. bei Erinnerungsaussagen, oder bei Bildanschauungen treten sie viel später auf. Es handelt sich also nicht um Epochen, welche die Gesamtpersönlichkeit in ihrer Entwicklung durchläuft (denn dann müßte diese auf einer bestimmten Altersstufe in jeder Beziehung einem der

Stadien angehören), sondern um Durchgangsstufen, welche die Bewältigung jeder intellektuellen Aufgabe durchmessen muß. Von dem Termin, in welchem normalerweise die Aufgabe an das Kind herantritt, und von ihrer Schwierigkeit hängt es ab, wann hier die Stadien erreicht werden. Darum kann ein Kind, das für ein bestimmtes Leistungsgebiet im „Aktionsstadium" ist, für ein anderes, schwereres oder leichteres, in einem tieferen oder höheren Intelligenzstadium sein.

Für die allgemeine Psychologie des Denkens ist das kategoriale Entwicklungsgesetz deswegen so bedeutsam, weil es zeigt, wie in allen Tätigkeiten, auch den einfachsten Auffassungsakten, bereits eine starke Aktivität und Spontaneität der Persönlichkeit beteiligt ist. Das Material an Sinneseindrücken, welches von der Außenwelt im Bewußtsein erzeugt wird, wird niemals einfach hingenommen, registriert und behalten; sondern es kann den Charakter einer Beobachtung, eines Gegenstandes der Aufmerksamkeit, einer Erinnerung, einer Darstellung erst gewinnen durch eine spontane Verarbeitung, die zugleich negativ und positiv vorgeht. Negativ, indem sie einen großen Teil des verfügbaren Materials abgleiten läßt; positiv, indem sie eine Auslese unter bestimmten Gesichtspunkten (eben den „Kategorien") vornimmt. Hemmung und Auslese gehen dabei — wenigstens bei freien Beobachtungen und Berichten — ohne willensmäßige Absicht vor sich; sie springen hervor aus einer inneren Zielstrebigkeit, die tiefer als alle Bewußtheit liegt und eben darum eine solch deutliche Gesetzmäßigkeit der Entwicklung zeigt.

Es ist ferner hervorzuheben, daß sich die Gesetzmäßigkeit vornehmlich auf die spontane Intelligenzbetätigung bezieht. Wird die Leistung stark durch äußere Einflüsse mitbestimmt, z. B. durch eindeutig abgegrenzte Aufgaben, durch die Fragen eines Verhörs usw., so wird das Stadiengesetz leicht verdunkelt. Wir müssen daher scheiden zwischen der Geneigtheit und der Möglichkeit, bestimmte Kategorien anzuwenden. Ein Kind, das bei einer gewissen Leistungsform (z. B. dem Bericht über ein Bild) spontan das „Substanzstadium" zeigt, ist nicht etwa unfähig, Aktionen, Qualitäten, Beziehungen aufzufassen, zu reproduzieren und darzustellen; es ist nur nicht von selbst innerlich dazu disponiert. Befragt man es über Vorgänge und Eigenschaften, so vermag es mancherlei Wissen darüber zu bekunden. Die anderen Kategorien sind also ebenfalls schon vorhanden und in gewissem Maße ver-

fügbar; aber sie entsprechen nicht der natürlichen und spontanen Einstellung des Kindes. Allerdings ist dann meist auch der intellektuelle Wert dieser entwicklungsfremden Kategorien ein geringerer; mit Recht sagt Meumann (*9*): „Bloß von der Umgebung eingetrichterte Anschauungsgesichtspunkte verraten ihren Ursprung durch die schablonenhafte und oft unzutreffende Anwendung von seiten des Kindes" (Bd. III, S. 404). — Aus alledem geht wieder hervor, daß die spontane I. aufs engste mit dem Faktor des Interesses und durch dessen Vermittlung mit Dispositionen des Gemüts verknüpft ist. Diese Zusammenhänge sind leider noch nicht genügend aufgehellt.

Man darf nach letzterem nicht erwarten, daß sich die I.-Stadien mit physikalischer Schärfe gegeneinander abheben, sie bedeuten nur ein Vorwiegen der geistigen Einstellung nach der einen oder anderen Richtung hin. Die bisherigen experimentellen Befunde ergaben meist eine relativ deutliche Abgrenzung des Substanzstadiums von den höheren Stadien, während innerhalb dieser die Grenzen zwischen der vorwiegenden Betonung der Handlung, der Eigenschaften und der Beziehungen mehr verfließen. Immerhin bleibt die Richtung der Entwicklung im genannten Sinne deutlich. Bei bestimmten intellektuellen Aufgaben ließ sich noch ein über die genannten Stadien hinausgehendes „Stadium der kritischen Reflexion" feststellen; dies führt uns aber zu höheren Altersstufen, von denen erst weiter unten die Rede sein wird.

Die empirische Feststellung der I.-Stadien ist entweder auf dem Wege der Beobachtung oder auf dem des Experiments möglich.

Das erste Verfahren ist namentlich bei kleinen Kindern eingeschlagen worden, bei denen man unmittelbar das spontane Verhalten der I. in ihrer Entwicklung verfolgen konnte. So erwähnten wir schon oben (S. 38), daß beim Sprechenlernen die Entwicklung des Wortschatzes dem kategorialen Gesetz zu folgen pflegt. Ähnliches ergab sich bei Feststellungen über die früheste Erinnerungsfähigkeit des Kindes. Die ersten Erinnerungen, die unsere Tochter um 1; 6 äußerte, bezogen sich ausschließlich auf Gegenständliches, besonders auf Personen; um 1; 9 begann sie sich an Vorgänge, namentlich an selbst getane Handlungen zu erinnern: endlich wurden Menschen und Handlungen in der Erinnerung bestimmt lokalisiert und damit die leichteste Art der Relation, die Ortsbeziehung, erworben. Erinnerungen an Merkmale, wie das Aussehen bestimmter Personen usw., traten erst um 2; 9 auf.

Bei experimentellen I.-Untersuchungen trat die Stadienfolge dort hervor, wo den Prüflingen ein Spielraum in der Auswahl der zu erwähnenden Gesichtspunkte gegeben war. Deshalb lieferte der spontane Bericht über eine gegenwärtige oder vergangene Wahrnehmung besonders deutliche Beweise für unser Entwicklungsgesetz. Bei Aussageversuchen (*40*) über ein eben gesehenes

Bauernstubenbild konnte ich erstmalig die Stadienfolge feststellen; und zwar zeigte sich in den Berichten der 7jährigen Kinder eine Alleinherrschaft oder unverhältnismäßiges Überwiegen unverbundener Gegenstandsbezeichnungen (Personen und Sachen); bei 10- und 11jährigen trat Nennung von Handlungen in größerem Umfange hinzu; bei 14jährigen und älteren begannen nähere Bestimmungen eine wachsende Rolle zu spielen, und zwar sowohl Beziehungen (namentlich örtliche) der Gegenstände untereinander wie auch Eigenschaften der Einzelgegenstände: Farbe, Form, Größe. Obgleich die Ortsangaben früher aufzutreten pflegen als die Eigenschaftsangaben, schien mir doch von einer nochmaligen deutlichen Trennung innerhalb des Relation-Merkmal-Stadiums keine Rede zu sein. (Die Altersangaben haben hier natürlich nur die Bedeutung von ungefähren Durchschnittswerten.) — Beispiele:

Ein 7jähriges Mädchen berichtete über das Bild folgendermaßen: „Ein Mann, eine Frau, eine Wiege, ein Bett, ein Knabe, ein Stuhl . . .“

Ein 10jähriges Mädchen: „Eine Familie. Der Vater sitzt am Tisch und ißt. Sein kleiner Sohn sitzt am Tisch. Das kleinste Kind liegt in der Wiege. Der Hund sitzt auf dem Boden und sieht auf den Tisch . . .“

Ein 14jähriger Knabe: „In der Mitte stand ein Tisch . . . Der Mann saß auf einem Stuhl, der Sohn auf einer Bank . . . An der linken Wand hing ein Bild und ein Kreuz. Die Hinterwand zeigte ein Fenster und zwei Bilder. Links war ein Bett, darunter ein Stiefelknecht . . .“

Ein 19jähriger Seminarist: „Eine Stube. Die Decke oben mit Holz getäfelt, wahrscheinlich Eichenholz. An der Wand ein Bild mit gelbem Rahmen . . . Ein Fenster mit Rouleaux, nicht ganz aufgerollt, schief, ein Bild darauf . . . Dann der Tisch vorn braun, gedrechselte Beine . . .“

Aussageuntersuchungen mit einem anderen Bild, das zu Deutungen und Reflexionen mehr Anlaß bot, als das Bauernstubenbild, führten Cohn und Dieffenbacher (*22*) dazu, außer jenen drei gleichfalls gefundenen Stadien, ein „kritisch-reflektierendes“ anzunehmen, das im höheren Jugendalter auftrat.

Entsprechendes ergab sich bei Anschauungsversuchen, d. h. solchen, bei denen der Gegenstand des Berichtes während des Berichtens vor Augen blieb. Cohn und Dieffenbacher fanden hier ihre vier eben genannten Stadien wieder; Schröbler (*34*), der Anschauungs- und Aussageversuche vergleichend anwandte, betont besonders die Trennung des Substanzstadiums von den höheren kategorialen Formen. Binet-Simon (*2, 3, 4*) fügten ihren bekannten Staffelsystemen der I.-Prüfung einen Bildbeschreibungsversuch ein und verlangten von 3jährigen Kindern „Aufzählungen“ der Personen und Gegenstände, von 7jährigen „Beschreibungen“ der Tätigkeiten und Zustände, von 12jährigen „Erklärungen“ der Zusammenhänge. — Man sieht ohne weiteres die Beziehung zu dem Entwicklungsgesetz[1]).

Die kategoriale Entwicklung besteht, so fanden wir, in der fortschreitenden geistigen Bewältigung der Außenwelt; sie vollzieht sich im wesentlichen innerhalb der Kindheit. Mit der Puber-

[1]) Daß auch bei freien Phantasieaufsätzen eine entsprechende Stadienfolge auftritt, hat Valentiner (*44*, S. 77) festgestellt.

tätszeit setzt nun eine neue Wendung der theoretischen Intelligenz ein, die von außen nach innen führt: der junge Mensch entdeckt sich selbst. Das eigene Ich ist nicht nur der Schauplatz und die Kampfstätte früher ungeahnter Gefühle, Sehnsüchte und Strebungen, sondern es wird auch zum Gegenstand einer geistigen Arbeit bisher unbekannter Art. Da handelt es sich zunächst um Selbstbeobachtung, um ein Sich-Rechenschaft-Geben von inneren Erlebnissen und Zuständen, die man bis dahin als selbstverständlich hingenommen hatte: dies erfordert Akte des Auffassens, Verstehens und Analysierens unter neuen isolierenden Gesichtspunkten. Es handelt sich ferner um eine Selbstbeurteilung der eigenen Motive und Eigenschaften: dazu müssen Tätigkeiten des Urteilens, des Deutens, des Kritisierens ins Spiel gesetzt, neuartige Wertmaßstäbe gefunden und angelegt werden. Es handelt sich drittens um Selbstgestaltung, die in die Zukunft weist: hier müssen die Hilfsmittel der praktischen I. — Vorwegnahme künftiger Ziele im Gedanken und Wahl der rechten Mittel zu ihrer Verwirklichung — aufs Psychische angewandt werden. Und schließlich handelt es sich um Selbstdarstellung: Umsetzung des neuerlebten Innendaseins in feste Begriffe und sprachliche Form, wobei Gedanke und Wort sich in ganz neuer Weise dem Schwebenden und Flüchtigen, dem Unanschaulichen und Ungreifbaren des seelischen Lebens anpassen müssen.

Grad und Art dieser Verinnerlichung der I.-Tätigkeit sind von Mensch zu Mensch sehr verschieden; doch bleibt wohl kein Jugendlicher ganz unberührt von dieser Wendung.

Die bloß subjektivistische Einstellung ist nun aber nichts Endgültiges. Vielmehr wendet sich die durch sie bereicherte und gewandelte I. wieder der Außenwelt zu, um ihr nun mit anderen Augen gegenüberzutreten. Drei neue Züge zeigt diese Welterfassung des Jugendlichen gegenüber der des Kindes.

Die im eigenen Ich entdeckte seelische Welt wird jetzt auch in andere Iche hineinprojiziert; der Akt der „Einfühlung" erhält seine intellektuelle Parallele in einer „Eindenkung". Die Erfassung, Deutung und Beurteilung fremder Individualitäten in ihren Gedankengängen und Willensmotiven, ihren Charakter- und Temperamentseigenschaften wird zu einer gern geübten geistigen Beschäftigung.

Es wird ferner das eigene Ich in seiner Stellungnahme zur Welt und die Welt in ihrer Beziehung zum Ich erfaßt: Re-

flexion, Werturteil und Kritik machen eine schnelle Entwicklung durch[1]); bald werden sie nicht nur auf Einzeltatsachen, sondern auch auf die grundsätzlichen Fragen des Daseins angewendet; die Umrisse einer Lebensanschauung werden gesucht.

Endlich aber wird bewußt der subjektive und objektive Anteil in dieser geistigen Arbeit getrennt. Die Welt in ihrer Loslösung vom Ich, in ihrer Unbekümmertheit um individuelle Ziele und Zwecke, in der Allgemeinheit und Notwendigkeit ihrer Gesetze wird zu einem eigenen Ziel des Denkens: die wissenschaftliche I. erwacht und damit diejenige Form, in der I. überhaupt ihre reinste und adäquateste Ausbildung erfährt.

Der Form nach aber ist diese I. der jugendlichen Reife durch ein gleichzeitiges schnelles Wachstum der analytischen und der synthetischen Kraft gekennzeichnet, unter Vorherrschaft der Synthese. So sehr der junge Mensch sich an der Entdeckung unendlicher neuer Einzelheiten berauscht, sich am Zerfasern, Kritisieren und Unterscheiden nicht genug tun kann: sein höchstes geistiges Streben geht doch auf die großen Grundfragen, die letzten Zusammenhänge und Verknüpfungen; er will in allem einzelnen die ewigen geistigen Ideen, im Zufall die Gesetzmäßigkeit, im Chaos des Lebens einen einheitlichen Sinn erfassen: das ist der philosophische Zug der jugendlichen I.

Es bedarf wohl kaum mehr der Betonung, daß mit diesen Charakteristiken nur bestimmte Neigungen und Richtungen angegeben werden sollen, in denen sich die I. im Kinde und Jugendlichen entwickelt. Über Grad und Art der tatsächlichen Entwicklung ist damit nichts gesagt. Der eine oder andere dieser Züge tritt bald nur in Andeutungen auf, bald in starker Ausprägung; bei manchen fehlt er ganz, bei anderen zeigt er sich in einseitiger Verzerrung. Nur Leitlinien einer entwicklungsgeschichtlichen Betrachtung der I. sollten in Obigem gegeben werden, nicht mehr.

---

[1]) Vgl. das oben erwähnte „kritisch-reflektierende Stadium", von dem Cohn und Dieffenbacher sprechen.

II. Teil.

Die Untersuchung der Intelligenz.

V. Kapitel.

Hauptgesichtspunkte der Methodik.

1. Allgemeines.

Aussagen über die I. der Menschen spielen in der Popularpsychologie des Alltags zu allen Zeiten eine gewaltige Rolle; in Schule, Haus und öffentlichem Leben wurden und werden fortwährend Urteile über Dummheit und Klugheit, über geistige Normalität und Abnormität (nach unten und oben) gefällt; es werden mit dogmatischer Selbstverständlichkeit Behauptungen ausgesprochen über die Abhängigkeit der I. von sozialer Stellung und Schule, von Vererbung und Geschlecht — aber alle diese Äußerungen entbehren der wissenschaftlichen Klärung und Sicherung und erweisen sich als durchaus unzureichend, wenn es gilt, auf Grund der I. der Menschen lebenswichtige Entscheidungen über ihr Schicksal zu treffen. Gewiß soll die natürliche Menschenkenntnis, die hierbei oft mitwirkt, nicht gering geschätzt werden; sie wird auch in Zukunft durch noch so exakte wissenschaftliche Methoden niemals überflüssig gemacht werden. Aber man darf sich nicht allein auf sie verlassen. Denn sie ist in sehr verschiedenem Maße auf die Menschen verteilt und steht mit ihrer Zuverlässigkeit leider oft in umgekehrtem Verhältnis zu der selbstgefälligen Sicherheit, mit der Urteile über die Fähigkeiten eines anderen abgegeben werden; sie ist ferner stets individualisierend und kann daher niemals allgemeine Einsichten, exakt vergleichende Feststellungen und Maßbestimmungen über die I. vermitteln; sie geht endlich rein intuitiv vor und vermag sich daher keine Rechenschaft zu geben über die einzelnen Gesichtspunkte und Faktoren, die bei der Beurteilung der I. in Betracht kommen.

Unsere Zeit konnte sich weniger als irgend eine andere hiermit zufrieden geben. Denn das Neue an ihr ist eine weitgehende Psycho-

logisierung des gesamten menschlichen Lebens. Die nichtpsychologischen Unterscheidungsmerkmale zwischen Mensch und Mensch, die früher ganz überwiegend alles Lebensschicksal bestimmten: die Verschiedenheiten des Besitzes und der Herkunft, treten zurück; an ihrer Stelle sollen die psychischen Fähigkeiten in früher unbekannter Weise entscheidend wirken. Fürsorgeeinrichtungen für Schwachbefähigte, Erziehungsmaßnahmen für Abnorme, Gliederung des Schulwesens nach Begabungsgraden und -richtungen, Auslese der Hochbefähigten, Berufsberatung auf Grund der geistigen Eignung, Beurteilung von Delinquenten auf Grund ihrer psychischen Beschaffenheit — das sind einige der Hauptmaßnahmen, die von einer neuen Ethik gefordert oder schon in Angriff genommen werden. Die Fähigkeiten werden in den Dienst der Allgemeinheit gestellt, und darum hat diese zu verlangen, daß sie in möglichst vollkommener Weise erkannt und behandelt werden.

Darum brauchen wir eine wissenschaftlich begründete Psychognostik und Psychotechnik der geistigen Fähigkeiten; und die allgemeine Befähigung, die wir Intelligenz nennen, bildet ein Hauptkapitel in diesem neuen Forschungsgebiet.

Die wissenschaftliche I.-Untersuchung ist wenige Jahrzehnte alt. Ihre ersten Objekte waren erwachsene, nicht-normale Menschen; zur Erforschung ihres Geisteszustandes bildeten die Psychiater eine große Reihe von Einzelmethoden der I.-Prüfung aus, die sie zum Teil auch schon zu ganzen Prüfungsserien verbanden (Rieger, Kraepelin, Ziehen, Sommer, Gregor, Bernstein, Rossolimo usw.).

Etwas später wandte sich das Interesse auch jugendlichen, nicht-normalen Individuen zu; denn Heilpädagogik, Hilfsschulwesen und Jugendgerichtsbarkeit — lauter neue Organisationen — bedurften einer genaueren Kenntnis der jugendlichen Individualitäten, als die bloße Alltagserfahrung sie gewähren konnte; hier gesellten sich zu den Psychiatern die Psychologen, wie Binet und Meumann, um geeignete Untersuchungsmethoden der unternormalen I. zu finden. Insbesondere hat das von Binet (*3*, *4*) geschaffene Staffelsystem hierfür eine weitgehende internationale Bedeutung erlangt.

Die Feststellung von Methoden für unternormal befähigte Kinder hatte aber zur Vorbedingung, daß man die normalen Maßstäbe kannte, an denen der Grad des Defektes zu messen war; und

so ergab sich die Notwendigkeit, auch die große Masse der normalen Kinder und Jugendlichen exakten I.-Untersuchungen zu unterziehen. Sehr bald aber gewann diese Arbeit ihr eigenes Gewicht; es zeigte sich, daß innerhalb der Normalität selbst gewaltige I.-Unterschiede bestanden, die in ihrer Verteilung, in ihrer Abhängigkeit von Alter, Geschlecht, Schule und sozialer Schichtung eine Fülle theoretisch und praktisch wichtiger Fragestellungen darboten.

Schließlich wurden die Grenzen der Normalität von der I.-Untersuchung auch nach der anderen Seite hin überschritten. Die Einrichtung von Begabungsschulen und das Bestreben, besonders geeignete junge Menschen bestimmten Berufen zuzuführen, verlangte Methoden der Auslese, die zum größten Teil von der Psychologie ausgebildet werden mußten; und so stellt die Erkennung der Höherbefähigten das letzte Glied in der Kette dieser Verfahrungsweisen dar.

So hat sich im Laufe von etwa zwei Jahrzehnten eine bereits sehr große Mannigfaltigkeit von Methoden entwickelt, über die nunmehr ein Überblick gegeben werden soll, sofern sie sich auf Kinder und Jugendliche erstrecken.

Wir beschränken uns also auf jenen Teil des Gebiets, der im wesentlichen von psychologischer und pädagogischer Seite her angebaut worden ist. Die psychiatrische Methodik, die insbesondere an Erwachsenen ausgebildet wurde, wird nur selten gestreift; sie hat auch ihre ganz besonderen Gesichtspunkte und Betrachtungsweisen, welche den hier zu behandelnden ferner liegen.

So reichhaltig das Methodenbild auch schon ist, so darf man nicht vergessen, daß wir es nur mit ersten Anfängen einer weit ausschauenden wissenschaftlichen Arbeit zu tun haben. Fast alles ist noch unfertig und verbesserungsbedürftig. Bewährtes steht noch hart neben weniger Zulänglichem; übertriebene Hoffnungen und voreilige Ansprüche fehlen ebensowenig wie verständnislose Befehdungen und unfruchtbare Kritteleien. Aber eben deshalb dürfte eine Gesamtübersicht erwünscht sein, die die positiven Leistungen der Ergebnisse in vollem Maße würdigt, sich jedoch zugleich über die vorläufigen und endgültigen Grenzen Rechenschaft gibt, die diesen Bestrebungen gesetzt sind.

Ihrer methodischen Beschaffenheit nach gliedern sich die I.-Untersuchungen nach drei Dimensionen: 1. in beobachtende und experimentierende Verfahrungsweisen, 2. in Forschungs- und

Prüfungsmethoden, 3. in Feststellungen über Einzelfunktion und über die Gesamtintelligenz eines Menschen. Der Besprechung dieser drei Gesichtspunkte sind dann noch einige allgemeine Erörterungen über Maßmethoden der I. anzufügen.

2. Experiment (Test) und Beobachtung.

Da sich die psychologische Methodik in einem gewissen Gegensatz zu der Unzulänglichkeit naiver Menschenbeurteilung entwickelt hat, ist es verständlich, daß sie zunächst das Hauptverfahren jener, nämlich die zwanglose Beobachtung des natürlichen Verhaltens, ablehnte und sie durch ein exaktes und kontrollierbares Experimentalverfahren, das der I.-Tests, zu ersetzen suchte.

Unter einem „Test" versteht man bekanntlich eine scharf umschriebene Aufgabestellung, die zur Bekundung einer bestimmten psychischen Beschaffenheit führen soll. Von dem psychologischen Laboratoriumsexperiment unterscheidet sich der Test dadurch, daß er in kurzer Zeit mit einfachen Hilfsmitteln einen für das Individuum kennzeichnenden Individualwert erzielen soll; er ist eine Art psychischer Stichprobe. Von den üblichen pädagogischen Prüfungen unterscheidet er sich dadurch, daß er nicht objektive Leistungserfolge als solche bewertet, nicht die Ergebnisse irgend einer Unterrichts- und Lernarbeit kontrollieren, sondern die den Leistungen zugrunde liegenden Fähigkeiten verraten soll; er ist nicht Kenntnis-, sondern Fähigkeitsprüfung.

Folgende Anforderungen werden an die Tests gestellt:

1. Der einzelne Test soll imstande sein, gewisse Seiten des psychischen Verhaltens zu isolieren, und zwar solche Seiten, die in der inneren Veranlagung des Menschen begründet sind. So gibt es Tests für die Ausdauer der Aufmerksamkeit, für die Merkfähigkeit für sinnvolle und sinnlose Stoffe, für die sprachliche, die logische, die optisch räumliche Kombinationsfähigkeit, für die Kritikfähigkeit usw. Nur ganz allmählich hat die Psychologie es gelernt, die äußeren Einflüsse von Kenntnis und Übung auszuschalten oder wenigstens zurückzudrängen und Tests zu ersinnen, bei deren Lösung wirklich die Fähigkeit als solche die Hauptrolle spielte. Es mußten tatsächlich neue Anforderungen sein, die an den Prüfling herantraten, damit es die wahre I. war und nicht Gedächtnis und Routine, die in der Testleistung zur Geltung kam.

2. Durch eine geschickte Verbindung mehrerer Tests soll eine möglichst vielseitige Prüfung der zur I. gehörigen Tätigkeiten gewährleistet werden. Über solche Testserien und Testsysteme wird weiter unten mehr zu sagen sein.

3. Der Test soll es ermöglichen, alle Prüflinge unter vergleichbare Bedingungen zu stellen. Dies geschieht entweder, indem alle Individuen mit genau demselben Verfahren, dessen Gang bis ins Einzelnste festgelegt ist, untersucht werden, oder dadurch, daß der Test in exakt kontrollierbaren, womöglich gradweise abstufbaren Abwandlungen vorgelegt wird.

4. Die Ergebnisse des Tests sollen ihrerseits quantitativ bewertbar sein, um eine weitgehende Vergleichung der Prüflinge zu ermöglichen. Die Bewertung geschieht entweder in der gröberen Form, daß der Test als gelöst oder verfehlt (+ oder —) verzeichnet wird: „Alternativtest", — oder in der feineren Form, daß jede Leistung eine Wertziffer erhält, die eine Rangordnung der Prüflinge erlaubt: „Stufentest". Die quantitativen Testergebnisse können dann wieder zu Resultantenwerten verbunden, zu Durchschnittswerten vereinigt, zu Korrelationsbestimmungen verwertet werden. So wird durch die Testprüfung eine regelrechte „Intelligenzmessung" möglich.

5. Das Testverfahren hat endlich den großen Vorzug der kurzen Dauer und vermag in einer oder wenigen Sitzungen über sonst unbekannte Menschen ein Urteil zu verschaffen, zu dem die bloße Beobachtungsmethode eine lange Bekanntschaft mit dem Individuum braucht.

Auf Grund solcher Erwägungen wurde in der angespannten Arbeit zweier Jahrzehnte, an der sich Psychologen, Lehrer und Ärzte der verschiedensten Kulturländer beteiligten, eine kaum übersehbare Mannigfaltigkeit von I.-Tests vorgeschlagen, erprobt, geeicht und zu praktischen Zwecken verwendet; und es ist eine wahrlich achtungswerte Leistung, auf die wir in dieser Hinsicht zurückblicken dürfen.

Aber sie führte zur Einseitigkeit. Die Vorzüge des Testprinzips veranlaßten manche, in ihm das allein seligmachende Verfahren zu erblicken; Meumann, Binet, Moede-Piorkowski und andere verlegten ihre Arbeit ausschließlich auf die Ausarbeitung und Verwertung von Tests und schrieben ihnen eine erschöpfende Leistungsfähigkeit für alle Feststellungen über die I. von Kindern

und Jugendlichen zu. Erst die letzten Jahre ließen uns die Unzulänglichkeit dieses Standpunktes erkennen und führten zu neuen methodischen Schritten.

Zunächst zeigte sich, daß die genannten Vorzüge zwar solche des Testprinzips sein mochten; daß aber der tatsächliche Stand unserer Testmethodik noch weit davon entfernt ist, jenem Ideal zu entsprechen. Wohl gibt es eine außerordentlich große Zahl möglicher Tests; aber erst langsam kommen wir dazu, die wirklich wertvollen Prüfmittel aus dieser Fülle herauszufinden. Wertvoll wird nämlich ein Test erst dann, wenn er einen hohen „Symptomwert" hat, d. h. als tatsächlicher Anzeiger für die I.-Beschaffenheit des Prüflings betrachtet werden kann, und wenn wir zugleich wissen, für welche Seite der I. er als Stichprobe gelten darf. Nicht jeder „Kombinationstest" prüft wirklich die Fähigkeit des „Kombinierens" oder diese allein; sprachliche Fertigkeit, Übung im Rätselraten, Interesserichtung bestimmter Art und manches andere mag mitwirken, um das Ergebnis zu erzielen. Und Gleiches gilt für jeden Test; eingehende psychologische Analysen sind nötig, um festzustellen, welche Seiten des Geisteslebens wirklich ins Spiel gesetzt werden; und solche Analysen fehlen uns zum Teil noch. Aus gleichem Grunde können auch die Zusammenstellungen von Tests für bestimmte Untersuchungsaufgaben noch nicht als endgültig befriedigende angesehen werden.

Außer diesen Unvollkommenheiten, die aus dem gegenwärtigen Stand der Testforschung hervorgehen und deren Überwindung durch fortschreitende Arbeit zu erhoffen ist, sind aber die endgültigen Grenzen der experimentellen Intelligenzuntersuchung zu beachten. Gerade das Testprinzip bringt es mit sich, daß der Test allein gar kein zureichendes Bild der I. zu geben vermag; denn er zeigt uns die Fähigkeit des Menschen, sofern sie reaktiv ist, nämlich auf scharf umrissene Einzelanforderungen des Augenblicks einzugehen vermag; zur I. gehört aber auch die Spontaneität, die von innen heraus erfolgende Bestimmung der Anlässe und Wege zur Betätigung der Denkfunktion. Die von uns im ersten Abschnitt aufgestellte, früher vernachlässigte Unterscheidung der reaktiven und spontanen Intelligenz (s. S. 18) gewinnt jetzt also auch eine grundsätzliche methodische Bedeutung.

Die Berücksichtigung der Spontaneität hat zunächst zu einem bemerkenswerten Fortschritt innerhalb der Testmethode selbst geführt. Wenn auch jeder Test eine Anforderung von außen her

stellt, so kann doch der Grad der Bindung, die er auferlegt, sehr verschieden sein. Die früheren Tests waren dadurch gekennzeichnet, daß sie reine gebundene Reaktionen verlangten, d. h.: es sollte eine Leistung ganz bestimmter eindeutiger Form vollzogen werden — so wenn die Merkfähigkeit für dargebotene Stoffe, die Aufmerksamkeitseinstellung auf gewisse zu durchstreichende Buchstaben, das Zusammensetzen von Figuren aus Bruchstücken untersucht wurde. Die neueren Tests erkennen aber in immer steigendem Maße die Spontaneität innerhalb der Reaktion an: es wird die allgemeine Anforderung zwar von außen gestellt — sonst wäre es ja kein Test; in der Ausführung selbst aber wird ein Spielraum für die spontane Auslese des Denkens gegeben; und gerade diese ist nun das eigentliche Problem des Tests geworden. Beispiele hierfür sind die Tests, die die Definition von Begriffen, die Beschreibung eines Bildes, die Bildung von Sätzen aus Stichworten, die Kritik einer mit Widersinnigkeiten durchsetzten Geschichte verlangen. Die größere Freiheit, die der Test dem Prüfling bietet, verringert zwar die Exaktheit im physikalischen Sinne, weil nun recht verschiedene seelische Faktoren mitschwingen können und die Ergebnisse schwerer meßbar und vergleichbar werden; wir gewinnen aber dafür, was diesen Nachteil mehr als aufwiegt, einen Einblick in die spontanen Triebkräfte und Richtungen der I.; und wir müssen uns zufrieden geben, wenn Feststellungs- und Wertungsmethoden eine gröbere Struktur und eine geringere Exaktheit erhalten. Mir scheint gerade in dieser Einbeziehung der Spontaneität eine der aussichtsvollsten Wendungen der neuesten Testforschung zu liegen.

Aber auch sie reicht nicht aus, um der Spontaneität die gebührende Beachtung zu sichern. Spontan ist hier zum Teil zwar die Art der Ausführung, nicht aber die Inangriffnahme der I.-Tätigkeit selbst. Erst dort, wo der Mensch sich unbefangen mit der Mannigfaltigkeit der Umweltreize auseinandersetzt, wo er selbst die Gelegenheiten zu intelligenten Handlungen aufsucht und aussucht, oder auch sie unbeachtet läßt oder vermeidet, wo er die ihm angemessensten Hilfsmittel und Richtungen des geistigen Tuns aus eigenem Ermessen bestimmt, kurz, wo die Intelligenz im Dienst des spontanen Interesses steht, gewinnt ihr Bild erst die Fülle der Farben und Linien. Diese Spontaneität aber kann überhaupt nicht mehr vom Prüfer in bestimmten Augenblicken durch bestimmte Anforderungen hervorgerufen werden; sie offenbart sich

nur dem Beobachter, der das natürliche Verhalten des Menschen mit liebevoller Sorglichkeit und Geduld studiert.

Aus solchen Gesichtspunkten heraus wird es verständlich, daß gerade aus den Kreisen der Praktiker (der Lehrer und Ärzte) immer wieder Mißtrauen gegen die nur-experimentelle I.-Untersuchung laut wurde. Aber andererseits bleiben doch die Bedenken, die wir oben gegen die ungeregelte „natürliche Menschenkenntnis" vorbrachten, in ihrer ganzen Schwere bestehen. Deshalb erwächst hier für die Psychologie die neue Aufgabe, die Methode der Beobachtung selbst und die Verwertung ihrer Ergebnisse exakt zu gestalten.

Die Beobachtungstätigkeit selbst muß jenen überlassen bleiben, die Gelegenheit haben, längere Zeit mit dem Kinde zusammen zu leben, also den Eltern, Lehrern, Ärzten, Fürsorgern usw.; der Psychologe aber hat die Vorbereitung dieser Beobachtungen, die Anleitung zu ihnen und die Bearbeitung ihrer Ergebnisse zu übernehmen. Was den laienhaften Beobachtern vor allem fehlt, ist die Leitung durch psychologische Gesichtspunkte; sie wissen nicht, worauf sie achten sollen; sie sind zu sehr gewohnt, die Handlungen und Leistungen der Kinder nach ihrem äußeren — moralischen oder didaktischen — Wert zu beurteilen, als daß sie die ihnen zugrunde liegenden seelischen Verhaltungsweisen zu erfassen vermöchten. Will man hier grundlegende Veränderungen erreichen, so müssen die Jugenderzieher ganz anders als bisher im psychologischen Beobachten geschult werden — dies ist die eine Aufgabe für den Psychologen[1]). Sodann aber muß der Psychologe für jede besondere Beobachtungsaufgabe die in Betracht kommenden psychologischen Gesichtspunkte darbieten, die vom Beobachter zu berücksichtigenden Fragen formulieren; denn: recht gefragt ist halb beantwortet. Dies geschieht durch psychographische Schemata, sogenannte Beobachtungsbogen[2]). Ferner kann der Psychologe den Beobachter anleiten, seine gesamten Wahrnehmungen und Teilurteile über die Intelligenz der Kinder zu einem Gesamturteil zu verdichten und danach die Kinder in eine I.-Rangordnung zu bringen: das ist die Methode der Intelligenzschätzung[3]). Er hat schließlich die so zustande gekommenen Befunde zu über-

[1]) Vgl. meinen kleinen Aufsatz: Erziehung zum psychol. Beobachten und Denken (*115*).

[2]) Ausführliches hierüber in Kap. XII.

[3]) Vgl. Kap. X, 2.

nehmen, um aus ihnen durch Vergleichung Typenbilder, Korrelationsbeziehungen und ähnliches abzuleiten.

Wird das Beobachtungsverfahren in dem angedeuteten Sinne ausgebaut, so verfügt die I.-Psychologie über ein System von Untersuchungsmethoden, die in ihrer Mannigfaltigkeit den verschiedenen Seiten des Problems gerecht werden und die in ihrer Verbindung ein wirklich umfassendes Bild der kindlichen I. zu geben vermögen. Das bloße Testverfahren wird dann nur noch dort erlaubt sein, wo aus äußeren Gründen ein sofortiges Urteil über Menschen gefällt werden muß, deren eingehende Beobachtung nicht möglich ist, oder wo sich im Massenversuch die Heranziehung so zahlreicher Beobachtungen von selbst verbietet. Aber die auf Grund der bloßen Testprüfung zu stellende I.-Diagnose wird stets nur einen vorläufigen Charakter tragen; der Test stellt für sich allein lediglich das „psychographische Minimum" dar. Das bloße Beobachtungsverfahren andererseits ist zu ungleichartig in der Ausführung von Mensch zu Mensch, um vergleichbare —, zu sehr bloß qualitativ, um meßbare —, und zu zeitraubend, um schnelle Auskünfte über kindliche I. zu geben; es vermag auch nicht, die Ausschaltung von Momenten, die nicht zur eigentlichen I. gehören, zureichend zu vollziehen. So ergänzen sich die beiden Methoden, die fälschlicherweise von der einen und von der anderen Seite als die alleinigen hingestellt werden. Und die Systematik ihrer Verbindung wird die künftige Hauptaufgabe der praktisch psychologischen Arbeit auf diesem Gebiet sein. Bereits sind in den jüngsten Begabungsauslesen vielversprechende Anfänge hierfür gemacht worden.

3. Forschungs- und Prüfungsmethoden.

Die Intelligenzforschung will Erkenntnisse über Wesen, Differenzierung und Bedingtheit der I. vermitteln. Hierzu bedarf sie einer Massengrundlage; denn nur, wenn sehr viele Individuen mit gleichem Verfahren untersucht werden, ist es möglich, sie unter bestimmten Gesichtspunkten (Alter, Schulklasse, soziale Schicht, Geschlecht, Schulleistungen usw.) zu gruppieren und Durchschnittswerte abzuleiten, aus denen die Beziehung der I. zu jenen Gesichtspunkten ersichtlich wird.

Deshalb ist die Massenstatistik ein wesentliches Hilfsmittel der I.-Forschung. Aber bei der großen Gefahr, die das bloße Zahlenwesen und der mit ihm verbundene Schein der Exaktheit mit sich bringt, muß hier besonders eindringlich zur Vorsicht ge-

mahnt werden; denn es ist gerade auf diesem Gebiet viel durch beliebige Rechnerei gesündigt worden.

Zweierlei ist nötig. Erstens müssen alle Vorsichtsmaßregeln des statistischen Verfahrens selbst berücksichtigt werden. Es dürfen nur solche Individuen zu Gruppen zusammengefaßt werden, die in bezug auf den zu untersuchenden Gesichtspunkt genügende Gleichartigkeit besitzen. Die geprüfte Gruppe muß als wirkliche Vertretung der betreffenden Kategorie von Individuen gelten können; d. h. die Gruppe darf an Gliedzahl nicht allzu klein und darf nicht einseitig zusammengesetzt sein. Der resultierende Wert für jede Gruppe ist entweder als Durchschnittswert (arithmetisches Mittel) oder als Zentralwert (Wert des nach der Rangordnung mittelsten Individuums der Gruppe) zu bestimmen. Außerdem aber muß auch der Grad der Streuung innerhalb der Gruppe, d. h. die mittlere Abweichung aller Einzelwerte von dem resultierenden Wert berücksichtigt werden.

Besondere Sorgsamkeit erfordern die Vorbereitungen von Massenuntersuchungen. Man sollte sie nie beginnen, ehe man sich über die anzuwendende Methode bis ins einzelnste Gewißheit verschafft, und ehe alle Mitwirkenden auf das genaueste eingeübt sind. Deshalb sind Vorversuche an einer geringeren Anzahl von Personen unbedingt erforderlich. Eine Hauptschwierigkeit bereitet der Umstand, daß der Psychologe meist nicht imstande ist, allein die gesamte Untersuchung vorzunehmen. Die große Zahl der erforderlichen Prüflinge kann ein einzelner überhaupt nicht oder nur innerhalb sehr langer Zeit bewältigen; oft aber ist die schnelle, ja womöglich gleichzeitige Erledigung aller dazugehörigen Einzelprüfungen geboten, um einen Gedankenaustausch der Versuchspersonen zu verhindern. Darum wird Arbeitsteilung nötig; mit der Heranziehung von Hilfsexperimentatoren kommt eine Ungleichmäßigkeit in die Prüfungsbedingungen; denn der Prüfer ist ein nicht unwichtiger Faktor in dem Reizgefüge, von dem der Ausfall eines psychologischen Versuchs abhängt. Diese Ungleichmäßigkeit läßt sich nicht ganz beseitigen, wohl aber auf ein Mindestmaß bringen durch eine peinlich genau ausgearbeitete und gründlich eingeübte Untersuchungsanweisung, die sich auf jedes zu sprechende Wort, auf jeden Handgriff, auf die Zeitdauer jeder Einzelphase bezieht. Das gleiche gilt auch für die statistische Verarbeitung des Massenmaterials; auch hier müssen alle Hilfskräfte nach genau demselben Wertungsverfahren arbeiten.

Man ersieht hieraus, daß die Beteiligung an psychologischen Massenuntersuchungen Schulung und Vorbereitung erfordert. Interesse und pädagogische Erfahrung allein reichen hierzu nicht aus. Ja die pädagogische Einstellung, die die berechtigte Neigung zum Individualisieren hat, muß gewaltsam zurückgedrängt werden, wenn man an psychologischen Massenuntersuchungen mitwirken will.

Wie verwickelte Organisationen unter Umständen nötig sind, um diesen Anforderungen der Massenuntersuchungen zu entsprechen, dafür bieten die Arbeiten von Minkus (*31*) in Breslau und die Begabungsauslese in Hamburg (*8*, I) (Kap. XII, 2) Beispiele.

Aber auch bei noch so sorgsamer Durchführung des Massenversuchs und statistischer Bearbeitung der Ergebnisse darf man nicht glauben, damit allein die Probleme erledigen zu können. Vielmehr muß die eigentliche psychologische Analyse dauernd Hand in Hand mit ihr gehen, damit man den wahren Sinn der Zahlenwerte richtig verstehe. Massenuntersuchungen können leicht dazu verleiten, daß man sich bei der Bearbeitung lediglich mit den schwarz auf weiß vorliegenden Befunden begnügt, mit den schriftlichen Testleistungen oder den Protokollen und Wertziffern der beteiligten Hilfskräfte. Wie jene Leistungen psychologisch zustande gekommen seien, bleibt dabei ganz unbekannt. In solchen Fällen ist es dringend nötig, daneben das gleiche Verfahren bei einer geringeren Anzahl von Kindern im Einzelversuch vorzunehmen, um die Kinder bei der Arbeit selbst beobachten, sie individuell befragen und auch die Versuchsanordnung individuell abwandeln zu können; dann treten erst die psychologischen Faktoren des Arbeitens hervor.

Von den verschiedenen Aufgaben, die mit Hilfe massenstatistischer Methoden bearbeitet werden, seien folgende genannt:

a) Die durchschnittliche gruppenmäßige Abhängigkeit der I. von biologischen, sozialen und schulischen Bedingungen.

b) Die Korrelation, d. h. das gemeinsame Variieren der einzelnen I.-Funktionen untereinander, sowie der Intelligenz mit der Schulrangordnung.

c) Die Verteilung der Intelligenzen innerhalb bestimmter Gruppen und in der Gesamtheit. Hierbei kann einerseits die Verteilung auf verschiedene I.-Typen, andererseits die relative Häufigkeit der verschiedenen I.-Grade studiert werden.

d) Die Eichung, d. h. die Feststellung der Angemessenheit

eines bestimmten Prüfungsmittels für die Angehörigen einer bestimmten Gruppe und die darauf gegründete Schaffung einer Maßeinheit für die betreffende Gruppe.

Die letztgenannte Aufgabe bildet die Überleitung von der I.-Forschung zur I.-Prüfung (IP.). Diese hat stets psychographische Absicht; sie will nicht Erkenntnisse genereller oder differentieller Natur, sondern die Kenntnis einzelner Individuen vermitteln, und verfolgt stets zugleich praktische Ziele. Das Ziel der IP. ist diagnostisch, wenn der gegenwärtige I.-Stand des Prüflings festgestellt werden soll; es ist zugleich prognostisch, wenn auch über die zu erwartende künftige Entwicklung ein Urteil abgegeben wird. Mit der Diagnostik begnügt man sich etwa, wenn in Frage steht, ob ein Kind zurzeit reif sei, in eine bestimmte Klasse aufgenommen zu werden, oder wenn beim Jugendgericht die Einsichtsfähigkeit eines jugendlichen Verbrechers geprüft wird. Meist allerdings sollen die IP. auch Prognosen geben, nämlich überall dort, wo sie bei einer Lebensbahnberatung mitwirken sollen. Sowohl die Bestimmungen über die Schulbahn (Überweisung zur Hilfsschule, Auslese für eine höhere Schulform) wie die Berufsberatungen haben ja weit ausschauende Bedeutung und gehen von der Voraussetzung aus, daß bestimmte Grade und Arten geistiger Fähigkeiten aller Wahrscheinlichkeit nach dauernd vorhanden sein werden.

Nun wurde bereits früher mehrfach darauf hingewiesen, daß die Entwicklung der I. nicht eine eindeutige ist, daß in ihr zuweilen Brüche und Wendungen vorkommen, die jeder Vorhersage spotten. Darum wird die Prognose der I. selbst dann keine volle Sicherheit gewähren, wenn die Methode zur Bestimmung des augenblicklichen I.-Zustandes ganz einwandfrei sein sollte. Und da auch in letzterer Hinsicht zurzeit noch keine Vollkommenheit erreicht ist, so wird man sich der Grenze unserer prognostischen Leistungsfähigkeit durchaus bewußt bleiben müssen.

Indessen hat man keinen Anlaß, deshalb die IP. zu verwerfen. Prognosen müssen gefällt werden, sobald man über das Schicksal von Jugendlichen Bestimmungen trifft; und die ganz laienhaften Vermutungen auf Grund bloßer Alltagserfahrung oder die rein schulmeisterlichen Voraussagen auf Grund der äußeren Schulleistung sind für sich jedenfalls noch viel unsicherer, als wenn außerdem noch als weitere Urteilsgrundlagen die I.-Prüfung und die geleitete Beobachtung herangezogen werden. Auch ist zu hoffen,

daß mit ständigem Ausbau des Verfahrens sein prognostischer Symptomwert immer mehr entwickelt werden kann.

Freilich, soviel lehrt auch diese Betrachtung: die IP. allein kann nicht die Verantwortung für lebenentscheidende Schulbahn- und Berufsberatungen tragen; wieder erweist sich die gemeinsame Benutzung aller Methoden des Experiments und der Beobachtung als erforderlich.

Eine trotz ihrer Selbstverständlichkeit nicht immer berücksichtigte Forderung besteht darin, daß man nur geeichte Prüfmittel zur IP. verwenden darf. Das heißt, man muß die Leistung des Prüflings auf die normale Leistungsfähigkeit der Gruppe beziehen, zu der er gehört; deshalb muß man solche Methoden benutzen, bei denen Lösung oder Nichtlösung für die I. der Gruppe kennzeichnend ist und für die der Prüfling nach der Art seiner Lösung eine bestimmte Stellung innerhalb seiner Gruppe erhält. Wie dies im einzelnen möglich ist, wird später bei den I.-Maßen zu besprechen sein. Benutzt man ungeeichte Prüfmittel, so kann man unter Umständen ein ganz schiefes Bild erhalten: zu schwere Aufgaben verwischen die Unterschiede zwischen Befähigten und Unbefähigten, weil hier wie dort lauter Versager auftreten müssen; zu leichte Aufgaben haben aus entgegengesetztem Grunde doch die gleiche nivellierende Wirkung, da hier die Älteren oder Befähigteren nicht mehr leisten als die Jüngeren oder Unbefähigteren. Hieraus ergibt sich, daß für jede Altersstufe jeweilig solche Prüfmittel gewählt werden müssen, bei denen die bestehenden Begabungsunterschiede durch deutliche Leistungsunterschiede gespiegelt werden.

Die IP. hat für praktische Zwecke schon jetzt eine sehr große Ausbreitung, namentlich in Amerika, erlangt; bei uns in Deutschland steht ihr, insbesondere im Hinblick auf die Einführung der nach Befähigungsart und -grad gegliederten Einheitsschule, ebenfalls eine weitgehende Verwendung in naher Zukunft bevor. Es muß nun dafür gesorgt werden, daß man hierbei nicht den Blick für den Ernst und die Schwierigkeit dieser Angelegenheit einbüße. Intelligenzprüfungen sind keine Spielereien, sind auch keine mechanischen Kunstgriffe, die jeder Dilettant sich in Kürze selbst aneignen kann, sondern sind höchst verantwortungsvolle wissenschaftliche Verfahrungsweisen, die ebenso eine wirkliche Vertrautheit mit der Technik des psychologischen Versuchs wie Schulung in psychologischer Analyse voraussetzen. Es wäre ein Unglück, wenn jeder Lehrer, der eine Vorlesung über Psychologie

gehört und etwas über IP. gelesen hat, sich nun für befugt hielte, sie praktisch anzuwenden. Die Psychologie muß auch als Gebiet des praktischen Lebens, ja gerade als solches wegen der damit verknüpften gewaltigen Verantwortung, Wissenschaft bleiben und darf nicht zum Tummelplatz dilettantischer Kurpfuscherei werden[1]).

Die kommende Ausdehnung der IP. wird daher die Ausbildung einer größeren Zahl von psychologisch geschulten Persönlichkeiten nötig machen, die von den Schulverwaltungen und Erziehungsanstalten für solche und verwandte Zwecke dauernd beschäftigt werden. Die von mir schon vor Jahren aufgestellte, von der Lehrerschaft zum Teil mißverstandene Forderung der Einführung von „Schulpsychologen" (*14*) wird sich durch den Zwang der Verhältnisse ganz von selbst durchsetzen, und bereits sind in manchen deutschen Städten schon Ansätze für diese Einrichtung vorhanden.

4. Untersuchung der Einzelfunktionen und der Gesamtintelligenz.

Wir definierten die I. als allgemeine geistige Anpassungsfähigkeit und drückten damit aus, daß sie sich über die verschiedensten seelischen Funktionen erstreckt, in denen der Mensch zu neuen Aufgaben Stellung zu nehmen hat. Deshalb kann die Prüfung der I. von den verschiedensten Seiten her erfolgen, unmittelbar, indem wir Denkleistungen verschiedener Art fordern: Tätigkeiten der Analyse und der Synthese, des Auffassens und Verstehens, des Urteilens und Kritisierens usw. — mittelbar, indem wir andere psychische Funktionen, z. B. die Aufmerksamkeit, die Merkfähigkeit, die Suggestibilität auf den in ihnen etwa enthaltenen I.-Anteil hin prüfen.

So wird es verständlich, daß gerade für die IP. eine unübersehbare Vielheit von Tests empfohlen und zum Teil auch angewandt worden ist. Sie sind von sehr verschiedenem Symptom-

[1]) Welch irrige Meinungen hier bestehen, geht aus Vorschlägen hervor, die Hauptmann Meyer (*87, 88*) vor mehreren Jahren machte. Er empfahl, daß bei der Rekrutenaushebung nicht nur körperliche Untersuchungen, sondern auch IP. stattfinden sollten, um ungeeignete Elemente von vornherein auszuschalten. Dieser Vorschlag ließ sich wohl hören; aber wenn er hinzufügte, daß die IP. hierbei von Unteroffizieren ausgeführt werden könnten, so muß dem auf das entschiedenste widersprochen werden. Psychologische Prüfungen sind denn doch etwas anderes als Feststellungen über Körperlänge und Gewicht, die von untergeordneten Organen ausgeführt werden können.

wert; aber jedenfalls haben wir es dieser Mannigfaltigkeit zu danken, daß unser Blick für die Vielgestaltigkeit der in der I. enthaltenen Teilfähigkeiten geschärft wurde. Die so möglich gewordene Analyse der I. ist die eine Vorbedingung jeder IP.; wir müssen wissen, was wir denn eigentlich mit jedem einzelnen Test prüfen. Eine Übersicht über die Hauptgruppen dieser Einzeltests ist in Kap. VI enthalten.

Es gibt so manche besondere Aufgaben der Forschung und Prüfung, bei denen wir uns auf bestimmte Teilseiten der Intelligenz (z. B. die Kombinationstätigkeit) bewußt zu beschränken haben; dann ist die Anwendung eines einzelnen Tests oder einer Verbindung einiger nahe verwandter Tests allenfalls zulässig. Niemals aber darf man glauben, mit dem Test oder den Tests eines einzelnen I.-Gebiets „die Intelligenz" zu treffen.

Dieser irrige Glaube hat unsere Arbeit zeitweilig gehindert. Man suchte nach dem I.-Test, nämlich nach einem einzelnen Prüfmittel, das blitzartig die Gesamtintelligenz des Individuums erhellen soll. So meinte Ebbinghaus (*60*), daß jede I.-Handlung letzten Endes im Kombinieren bestehe, und sah in seiner Testlückenergänzungsmethode die geeignete Stichprobe für die Kombinations-, und damit allgemein für die geistige Leistungsfähigkeit. Manche neuere Psychologen neigen dazu, einen allzu einseitigen Wert auf die Dreiwortmethode zu legen. Gregor (*64*, *65*, *65*a) arbeitet mit dem Definitionsverfahren usw. Vorgefaßte Theorien über das Wesen der I., zuweilen auch der Zufall, der den Forscher gerade auf den einen oder anderen Test führte, sind hierbei oft bestimmend gewesen. Gerechtfertigt wäre ein solches Verfahren nur dann, wenn man sicher wäre, mit einem solchen „Universaltest" alle wesentlichen Teilfunktionen der I. implizite mitzuprüfen. Diese Voraussetzung aber trifft nicht zu.

Gewiß ist die I. eine „allgemeine" Leistungsfähigkeit in dem Sinne, daß sie dem gesamten geistigen Verhalten des Individuums eine bestimmte gemeinsame Färbung verleiht. Aber wir betonten schon von vornherein[1]), daß man diesen Gedanken nicht bis zu einer mechanischen Gleichgradigkeit aller Teilfunktionen auf den verschiedenen Gebieten übertreiben dürfe. Dem widerspricht schon der Umstand, daß an jeder Einzelleistung nicht nur der I.-Grad, sondern auch der qualitative I.-Typ des Prüflings und daneben noch soundsoviele andere psychische Gebiete beteiligt sind.

[1]) Vgl. Kap. I, S. 3.

Wollen wir also wirklich ein Gesamturteil über die I. eines Menschen erhalten, so müssen wir sie von den verschiedenen Seiten her abzustecken suchen. Es müssen also aus der unübersehbaren Fülle möglicher Tests Testserien hergestellt werden — in genügender Mannigfaltigkeit, um die verschiedenen intellektuellen Teilfunktionen ins Spiel treten zu lassen, und mit möglichst hohem und sicherem Symptomwert der einzelnen Tests für das betreffende Teilgebiet. Die Schwierigkeit besteht darin, die Forderung der Vielseitigkeit und Gründlichkeit mit der anderen eines möglichst geringen Zeit- und Kraftaufwandes für Prüfling und Prüfer zu verbinden; eben darum hängt soviel von einer weisen Auswahl und Eichung der Tests ab; das willkürliche Aneinanderreihen beliebiger Prüfungsweisen, wie wir es noch oft finden, ist nicht angängig.

Über die wichtigsten bisher angewandten Testverbindungen berichten die Kapitel VII und XII.

Ein weiteres Bestreben geht dann dahin, aus dem so gewonnenen Mosaikbild eine einheitliche Gesamtbewertung der I. abzuleiten. Hierbei wird, da ja nicht alle Teilleistungen auf gleicher Höhe stehen, eine Art Ausgleichs-(Kompensations-)Verfahren angewandt werden müssen. Dies geschieht schon immer im Alltagsleben, wenn auch in unkritischer Weise; wir nennen A klüger als B, obwohl wir wissen, daß B in manchen Einzelfunktionen A übertreffen mag; ebenso ist auch in der Schule bei der Herstellung des Gesamtzeugnisses eines Schülers das System des Ausgleichs ausgebildet. Der Grundsatz ins Psychologische übertragen bedeutet: Zwei Individuen können, ohne in allen Punkten gleichartige I.-Leistungen zu zeigen, dennoch im ganzen gleichwertige I. besitzen; und es muß sich dieser Gesamtwert als Resultante aus den verschieden hohen Einzelleistungen jedes Prüflings auch zahlenmäßig ableiten lassen.

Genau die entsprechenden Gesichtspunkte gelten auch für das nicht-experimentelle Verfahren der I.-Untersuchung. Denn auch die Beobachtungsmethode ist zuweilen unerlaubt verengert worden in dem Glauben, daß es irgend ein Universalsymptom gäbe, aus dem man die I. ablesen könne (so wie etwa der Graphologe den Charakter lediglich aus der Handschrift zu erdeuten sich vermißt). Für den einen Lehrer gilt die Fähigkeit des Kindes in der Lösung angewandt mathematischer Aufgaben als eindeutiges Kennzeichen, der andere verläßt sich auf das Verständnis für geschichtliche

Zusammenhänge, ein dritter auf das Verhalten beim Spielen usw. Auch hier Einseitigkeit, Willkür und größte Buntscheckigkeit[1]).

Demgegenüber ist es Aufgabe des Psychologen, durch die Fragestellung des Beobachtungsbogens für eine wirkliche Vielseitigkeit der Beobachtung zu sorgen, derart, daß die verschiedenen Seiten der I. auch hier gleichmäßig zu ihrem Rechte kommen.

Und wiederum muß der Versuch gemacht werden, über das Mosaikbild hinaus zu einer einheitlichen Gesamtwertung zu kommen, sei es durch Einordnung des Kindes in einen I.-Typ, sei es durch Schätzung des I.-Grades und darauf gegründete Einreihung des Kindes in eine I.-Rangordnung, sei es endlich — mehr künstlerisch-intuitiv — durch Abfassung einer die Persönlichkeit des Kindes schildernden Charakteristik.

5. Intelligenzmaße.

Soweit die I. als eine einheitliche seelische Beschaffenheit angesehen wird, die bei jedem Individuum in einem bestimmten Grade vorhanden ist, muß es grundsätzlich möglich sein, diesen Grad durch einen Maßwert zu bezeichnen.

Solche Maßwerte sind tatsächlich in großer Anzahl und in bunter Mannigfaltigkeit vorgeschlagen worden. Sie sind nach folgenden Gesichtspunkten einzuteilen:

1. Relative und absolute I.-Maße. Das relative Maß will dem Prüfling lediglich einen bestimmten geistigen Platz innerhalb einer willkürlich gebildeten, in sich homogenen Gruppe anweisen, die gleichzeitig nach denselben Methoden geprüft worden ist. Dies geschieht, wenn man die Kinder einer gewöhnlichen Schulklasse oder die für eine Begabungsklasse in Aussicht genommenen Kinder prüft und untereinander vergleicht. Das absolute I.-Maß will dagegen die Stelle angeben, die dem Prüfling innerhalb seiner naturgegebenen Gruppe zukommt; es setzt also voraus, daß die Prüfungsmethode für jede Gruppe geeicht ist, daß sie einen Normalkanon erhält, der nun als Maßstab an jedes einzelne Individuum angelegt werden kann.

[1]) Sehr lehrreich ist in dieser Hinsicht eine Erhebung Binets (*50*). Er hatte eine Reihe von Volksschullehrern befragt, nach welchen Symptomen sie die I. ihrer Schüler zu beurteilen suchten; darauf wurden außer den schon oben benannten Kriterien u. a. auch folgende benannt: Gesichtsausdruck, Rechtschreiben, ausdrucksvolles Lesen.

2. Augenblicks- und Dauermaße. Das Augenblicksmaß gibt die jeweilige I.-Höhe an, welche das Kind in einem bestimmten Lebensjahr besitzt. Nun liegt uns aber daran, über diese augenblickliche Zuständlichkeit hinaus die konstitutionelle I. zu kennen, d. h. die tiefinnerste Beschaffenheit der I.-Anlage, die unabhängig vom Alter dauernd das Individuum durch das Leben begleitet. Erst ein solches Maß, dem wir Dauerwert zuschreiben dürfen, würde für das einzelne Individuum Prognosen gestatten und pädagogische Maßnahmen für die Zukunft unterstützen. Es würde auch durch seine Unabhängigkeit vom Alter weitgehende Vergleichungen erlauben, z. B. solche der Geschlechter, der Stände, der Nationen, der Schwachsinnsformen usw.

Selbstverständlich wird der Dauerwert der I. niemals — selbst wenn er unter den denkbar vollkommensten Bedingungen der Methodik erzielt wird — auch nur annähernd die Bedeutung physikalischer „Konstanten" gewinnen. Dies widerspricht dem Wesen der geistigen Eigenschaft, die als Bestandteil einer lebendigen Persönlichkeit stets auch von deren irrationalen und daher unvoraussagbaren Wandlungen mitergriffen wird; und es widerspricht der Erkenntnisfähigkeit des Psychologen, der es nicht wie der Physiker mit einem einfachen, starren und toten Gebilde zu tun hat. Ein Dauerwert der Intelligenz — wie wir ihn später z. B. im I.-Quotienten kennenlernen werden — hat daher niemals mehr als Wahrscheinlichkeitscharakter. Die Geschichte duldet eben keine wirkliche Prophetie — auch nicht die Geschichte des geistigen Individuums.

Die absoluten Maße gehen aus von dem Lebensalter des Kindes und stellen als Kanon diejenigen Leistungen auf, die von der betreffenden Altersstufe überhaupt normalerweise bewältigt werden können. Es war dies die bahnbrechende Idee Binets (*3*), welche allen weiteren Arbeiten auf diesem Gebiet zugrunde liegt: für jede Jahresstaffel der Kindheit ein System von Tests zu finden, welches eben für Leistungsgrad und -art dieser Altersstufe kennzeichnend ist.

Die Tests werden von Binet und fast allen seinen Nachfolgern als „Alternativtests" verwandt: es gibt bei ihnen nur die zwei Möglichkeiten der Lösung oder Verfehlung; eine Abstufung der Leistung ist beim einzelnen Test nicht meßbar. Lediglich durch Zählung der gelösten Tests werden die I.-Maße hergestellt. Diese Maße selbst, deren wichtigste das I.-Alter und der I.-Quotient sind, können nur im Zusammenhange des Binet-Simonschen Staffel-

systems besprochen werden; es sei daher auf Kap. VII, 3 verwiesen.

Dagegen müssen hier noch verschiedene Grundsätze der relativen I.-Messung erörtert werden, da an verschiedenen Stellen der weiteren Arbeit auf sie zurückzugreifen sein wird.

a) Rangplatz- und Korrelationsbestimmung. Das relative Maß soll dem Einzelindividuum einen bestimmten Platz innerhalb einer Gruppe anweisen. Voraussetzung ist hierfür erstens, daß die Glieder der Gruppe, abgesehen von ihrer I., einigermaßen gleichwertig sind (z. B. in bezug auf Alter, soziale Schicht, Muttersprache), zweitens, daß bei allen das gleiche Verfahren der I.-Festsetzung (sei es Schätzung oder Experimentalprüfung) angewandt wird.

Die unmittelbare Herstellung einer Rangreihe findet statt bei der I.-Schätzung, bei welcher nach dem allgemeinen Eindruck des Schätzers das Mehr oder Minder der I. (aber nicht die Größe des Unterschieds) in der Anordnung der Kinder zum Ausdruck gebracht wird. Über die besonderen methodischen Bedingungen der I.-Schätzung wird später (Kap. X, 2) zu sprechen sein.

Um aus der Testprüfung ein relatives I.-Maß für den Prüfling zu gewinnen, sind mehrere Wege gangbar.

Schon der einzelne Test kann eine Abstufung ermöglichen, sofern die Aufgabe, die er stellt, in abstufbarer Vollständigkeit und Richtigkeit gelöst werden kann. Ausgenommen sind hier also nur die oben erwähnten „Alternativtests". Alle anderen Tests erlauben eine Messung: bald ist es die Anzahl der Teilaufgaben, die gelöst werden (z. B. die Zahl der in einem Lückentext ergänzten Wörter), bald die Zahl der gemachten Fehler, bald die für eine vorgeschriebene Arbeit verwandte Zeit usw.

Ist die Gruppe lediglich mit einem einzigen Test geprüft, so kann die so gewonnene Maßzahl ohne weitere Umformung benutzt werden. Man gibt dann etwa an: bei einem Lückentext, bei welchem von der Gruppe durchschnittlich 30 Lücken richtig ausgefüllt wurden, hat X 25, Y 32, Z 40 Lücken ausgefüllt.

Will man aber die Leistungen eines Prüflings in verschiedenen Tests miteinander vergleichen oder zu einem Index der Gesamt-I. kombinieren, so ist vorher nötig, die Maßzahlen, die ja bei jedem Test etwas anderes bedeuten, auf einen gemeinsamen Nenner zu bringen. Als eine Möglichkeit zur Herstellung eines

solchen „gemeinschaftlichen Maßes" bietet sich nun die Rangmethode dar.

Für jeden einzelnen Test wird auf Grund der Maßzahlen eine Rangordnung der Prüflinge hergestellt. Prüflinge mit gleichen Maßzahlen erhalten gleiche Rangnummern, die der arithmetischen Mitte der von ihnen besetzten Rangplätze entsprechen. Für viele Fälle (z. B. beim Vergleich von verschieden starken Gruppen) empfiehlt es sich, die Rangplätze in Prozentnummern (%-Nr.) umzurechnen, die angeben, welchen Platz das Kind erhielte, wenn die Zahl der Prüflinge 100 betrüge. Dann bedeutet stets %-Nr. 1 den besten, %-Nr. 50½ den mittleren, %-Nr. 100 den schlechtesten Platz; durch die %-Nr. 26 und 75 wird die mittlere Hälfte aller Prüflinge abgegrenzt.

Sind für eine Gruppe von Prüflingen mehrere Rangordnungen hergestellt — sei es durch Schätzung und Prüfung, oder durch Prüfung in mehreren gesondert gewerteten Tests —, so kann die weitere Berechnung zwei Wege gehen.

Die Methode der Kombination verbindet mehrere Rangordnungen zu einer einheitlichen Gesamtrangreihe. Hat z. B. jedes Kind bei der Prüfung mit drei verschiedenen Tests drei verschiedene Rangplätze erhalten, so werden diese addiert; auf Grund der Summenwerte werden die Kinder in eine neue Rangordnung gebracht, die demnach ein gemeinsamer Ausdruck für die drei Leistungsformen ist. Ähnliches ist bei I.-Schätzungen möglich, wenn die verschiedenen in einer Klasse unterrichtenden Fachlehrer unabhängig voneinander die Schüler in eine Schätzungsreihe gebracht hatten. Solche „kombinierten Rangreihen" haben den Vorteil, daß sie etwaige Einseitigkeiten der Einzelreihen ausgleichen und daher dem „Allgemein"-Charakter der I. besser gerecht werden.

Die Methode der Rangkorrelation dient der Vergleichung zweier Rangreihen und stellt den Grad ihrer Übereinstimmung fest. Dieser Grad wird gemessen durch den „Korrelationskoeffizienten", der zwischen $+1$ und -1 schwanken kann. Beträgt er $+1$, so heißt dies, daß beide Reihen vollkommen übereinstimmen. Beträgt er -1, so laufen beide Reihen genau entgegengesetzt, d. h. die Individuen mit den besten Leistungen der einen Art zeigen die tiefsten Leistungswerte in der anderen Funktion und umgekehrt. Ist der Korrelationskoeffizient $= 0$, so besteht zwischen beiden untersuchten Gebieten keine Beziehung. Die Höhe des Dezimalbruchs zwischen 0 und $+1$ gibt also den Grad der positiven

Ähnlichkeit beider Rangreihen und damit den Grad des Zusammenhangs beider untersuchter Funktionen an.

Es gibt eine ganze Reihe von Berechnungsmethoden des Korrelationskoeffizienten[1]); an allen späteren Stellen dieses Buches ist, wo nichts anderes bemerkt wird, die einfachste Formel von Bravais und Pearson zugrunde gelegt.

Die Zahl der untersuchten (geschätzten, geprüften) Individuen sei n; irgend ein Individuum habe in der einen Rangreihe den Platz Nr. x, in der anderen Nr. y. Nun wird für jedes Individuum die Differenz (Nr. x — Nr. y) gebildet und quadriert, ferner werden die so entstehenden Quadratwerte aller Prüflinge addiert, so daß der Summenwert $\Sigma\,(\text{Nr. } x - \text{Nr. } y)^2$ entsteht. Dann wird der Korrelationskoeffizient durch folgende Formel ausgedrückt:

$$\rho = 1 - \frac{6\,\Sigma\,(\text{Nr. } x - \text{Nr. } y)^2}{n\,(n^2-1)}.$$

Das Maß der Zuverlässigkeit, die dem berechneten Wert zukommt, hängt von der Höhe dieses Wertes, wie von der Anzahl der untersuchten Individuen ab. Zur Berechnung des Zuverlässigkeitswertes dient der „wahrscheinliche Fehler“[2]):

$$wF(\rho) = -\,0{,}706\,\frac{1-\rho^2}{\sqrt{n}}.$$

Nur wenn ρ mindestens das Dreifache des wF beträgt, ist der Korrelation eine reelle Bedeutung zuzumessen. Ein konkretes Beispiel für die Berechnung der Rangkorrelation wird im Anhang gegeben.

b) Qualitätsziffern. Neben dem Rangverfahren gibt es noch eine andere Möglichkeit, um ein gemeinschaftliches Maß für verschiedene Tests zu gewinnen, nämlich die Benutzung von Qualitätsziffern (Prädikaten oder Punkten). Es werden dann gewisse Hauptqualitäten der Lösungen unterschieden, die auf alle geprüften Tests in gleicher Stufenzahl angewandt werden. Am praktischsten ist hier die Benutzung der auch in der Schule allgemein üblichen fünfgradigen Leiter: 1 = sehr gut, 2 = gut, 3 = genügend, 4 = mangelhaft, 5 = ungenügend.

Die vorher gewonnenen Maßzahlen für den einzelnen Test müssen hierfür in Qualitätsziffern übertragen werden, z. B.:

---

[1]) Gesamtdarstellungen der Korrelationsmethoden bei Betz (*47*) und Stern (*12*). Neuerdings hat Deuchler (*59*) eine veränderte Methode ausgebildet, die insbesondere für die Errechnung von Zeugniskorrelationen und verwandte Aufgaben gute Dienste zu leisten vermag. Auch auf das neue Rechnungsverfahren, das Lipmann (*198*) zur Bestimmung psychologischer Koordinationen ausgebildet hat, sei hingewiesen.

[2]) Diese Formel ist in der ersten Auflage dieses Buches, sowie in meiner „Differentiellen Psychologie“ mit einem Druckfehler wiedergegeben worden. Im Zähler darf kein Wurzelzeichen stehen.

0 Fehler = Prädikat 1, 1—3 Fehler = Prädikat 2, 4—8 Fehler = Prädikat 3 usw. Hierbei ist eine gewisse Willkürlichkeit nicht zu vermeiden; sie ist dadurch auf ein Mindestmaß zurückzuführen, daß man die Häufigkeitsverteilung der Prädikate möglichst dem normalen Verteilungsgesetz anzunähern sucht (s. Kap. VIII), nach welchem auf die Prädikate 1 und 5 ungefähr je 3%, auf die Prädikate 2 und 4 je 22%, auf das Prädikat 3 50% aller Fälle kommen[1]).

Die Vorzüge des Verfahrens bestehen im folgenden:

1. Man ist in der Lage, auch qualitative Eigenschaften der Leistungen, die sich in den direkten Maßzahlen nicht ausdrücken lassen (z. B. besonders geschickte Formulierung einer Kritik, Originalität einer Lösung), bei der Prädikatgebung mit zu berücksichtigen.

2. Das Verfahren ist insbesondere dem Pädagogen geläufig; die Ziffernergebnisse haben für ihn einen unmittelbar verständlichen Sinn.

3. Die Vergleichbarkeit und Vereinbarkeit der Ergebnisse verschiedener Tests ist in weitem Umfange gewährleistet.

Verbindet man bei jedem geprüften Kind sämtliche von ihm in n Einzeltests erarbeiteten Teilprädikate, so erhält es durch diese Gesamtziffer einen eindeutigen Platz zwischen der denkbar besten Leistung ($n \cdot 1$) und der denkbar schlechtesten Leistung ($n \cdot 5$). Diese Ziffer kann entweder als solche benutzt oder in einen Rangplatz bzw. Prozentrangplatz oder ein Gesamtprädikat ungewandelt werden. Eine Anwendung dieser Methode wird uns bei Besprechung der Hamburger Begabungsprüfungen begegnen.

Während die Prädikatziffern mit aufsteigender Zahl eine Verschlechterung der Leistung bedeuten, nehmen die „Punktziffern" gleichsinnig mit der Leistungsgüte zu. Man verwendet sie hauptsächlich bei einfachen und kurzen Tests, bei denen man über die bloße und grobe Alternativwertung „richtig" oder „falsch" zu einer etwas feineren Bewertung hinausgelangen möchte. Es wird dann auch das Gewicht oder die Sicherheit der Lösung eingeschätzt und danach die Punkte verteilt. So beziffert z. B. Piorkowski (*90*) bei der Dreiwortmethode die Erledigung jeder Stichwortgruppe folgendermaßen: Fehlende oder sinnlose Lösung = 0, teilweise Lösung = ½, richtige = 1, besonders gute und originelle = 1½.

[1]) Freilich verhindern die äußeren Verhältnisse bei manchen Tests nicht selten eine diesem Gesetz sich annähernde Verteilung der Prädikate.

Auch hier geht es nicht ganz ohne Willkür ab, sowohl bei der Festsetzung der Punktwerte wie bei der Anwendung. Aber wie die Praxis zeigt, ist diese Willkür das geringere Übel gegenüber dem Zwang, alle Lösungen in eine der beiden Kategorien „richtig" oder „falsch" ohne weitere Differenzierung zu pressen. Was vielleicht bei der Untersuchung seelischer Elementarfunktionen noch angeht, ist bei der I. nicht mehr möglich. Hier spielen die Übergänge und Zwischenstufen zwischen völligem Versagen und richtiger Lösung eine ganz bedeutende Rolle, hier ist auch innerhalb dessen, was als „richtig" angenommen werden kann, noch eine merkliche Abstufung in der Güte der Lösung vorhanden, so daß man sich immer wieder zur Einführung solcher Gewichtsbestimmungen veranlaßt sieht. Es wird hier vor allem darauf ankommen, Maß zu halten, und bei den verschiedenen Tests möglichst gleichmäßige und einfache Bestimmungen für die Qualitätsbezifferungen einzuführen.

Zu einem durchgearbeiteten System wurde die Punktmethode von Yerkes (*125*) erhoben, worüber S. 156/7 ausführlicher berichtet wird.

c) Seltenheitswert und Intelligenzzensur. Ein neuer Weg, der Güte einer I.-Leistung mit exakt zahlenmäßigen Hilfsmitteln beizukommen, besteht in der Berechnung ihres Seltenheitswertes.

Zuerst hat wohl W. Peters (*33*) in seinen Vererbungsuntersuchungen diesen Gesichtspunkt eingeführt (S. 356 und 357). Bei Versuchen mit der Dreiwortmethode führt er aus: „Wenn eine Aufgabe von allen 190 Knaben gelöst wurde, die an diesen Versuchen teilnahmen, so ist ihre Schwierigkeit offenbar eine geringere als die einer anderen Aufgabe, die nur von 95, also der Hälfte der Schüler, gelöst werden konnte. Wenn wir die Lösung der ersten Aufgabe mit 1 werten, so müssen wir die letztere doppelt so hoch, also mit 2 werten. Eine Aufgabe, die von $^3/_4$ der Schüler richtig gelöst wurde, ist demnach mit 1,33 zu bewerten, eine Aufgabe, die von 80% der Schüler gelöst wurde, mit 1,67 usw.".

Peters berechnet also den Schwierigkeitswert durch Umkehrung des Bruches, der die prozentuale Häufigkeit der Lösungen angibt. Der geringste Wert („von allen gelöst") ist $= 1$, der höchste („von keinem gelöst") $= \infty$.

Später ist Otto Schultze (*107*) unabhängig von Peters auf den gleichen Gedanken gekommen und hat ihn zur Gewinnung eines neuen I.-Maßes, der Intelligenzzensur (IZ.), verwertet, das an die Stelle des Binetschen „I.-Alters" treten soll. Er

rechnet anders als Peters; der Seltenheitswert (Schultze selbst braucht diesen Ausdruck nicht) ergänzt den prozentuellen Häufigkeitswert der Testlösungen H für das betreffende Alter zu 100%. Ein Test, der nur von 30% der 7jährigen gelöst zu werden pflegt, hat für die Kinder dieses Alters den Seltenheitswert 70%.

Nach Schultze ist demnach der geringste Seltenheitswert („von allen gelöst") = 0, der höchste („von keinem gelöst") = 100. Das Maß scheint bequemer und übersichtlicher als das von Peters zu sein; denn es besagt unmittelbar, einem wie großen Prozentsatz der Prüflinge der Test zu schwer war.

Ist nun ein Kind in einer größeren Anzahl von Tests (z. B. 25) verschiedener Schwierigkeit geprüft worden, für welche die normalen Lösungsprozente dieser Altersstufe bekannt sind, so kann man nach Schultze die Seltenheitswerte der vom Prüfling gelösten Tests vereinigen und zu den durchschnittlichen Seltenheitswerten aller vorgelegten Tests in Beziehung setzen. Diese Ziffer, welche die I.-Höhe des Kindes im Vergleich mit der durchschnittlichen I.-Höhe seiner Altersgenossen ausdrückt, ist die „Intelligenzzensur". Ihre Formel lautet: $IZ. = \frac{100 - H^b}{100 - H^p}$, wo H^b die Lösungshäufigkeit bei denjenigen Tests bedeutet, die der Prüfling seinerseits gelöst hat. Um den falschen Schein einer übertriebenen Exaktheit zu vermeiden, schlägt Schultze vor, die IZ. nur in Zehnteln auszudrücken. Die IZ. 0,2 würde also bedeuten: die Seltenheitswerte der gelösten Tests betragen $^1/_5$ der Seltenheitswerte aller geprüften Tests.

Die Einführung des Seltenheitswertes erscheint zunächst sehr einleuchtend; bei näherer Prüfung stellen sich jedoch der Durchführung des neuen Maßes ganz bedeutende Schwierigkeiten entgegen. Es ist überhaupt nur anwendbar, um die Stellung eines Kindes unter Kindern gleichen Alters, gleicher Bildungsstufe und auch ungefähr normaler Fähigkeit usw. zu charakterisieren, versagt aber bei der so wichtigen Aufgabe, Kinder verschiedener Gruppen miteinander zu vergleichen, oder die I.-Höhe eines Kindes durch verschiedene Altersstufen zu verfolgen. Denn wir haben gar keine Gewähr dafür, ob etwa die IZ. 0,3 bei 7jährigen dieselbe relative Bedeutung habe wie bei 9jährigen, da ja die Ziffer durchaus abhängt von der Auswahl der Tests, die hier und dort angewandt worden sind.

Wollte man hier Gleichmäßigkeit erzielen, so müßten die Testserien, mit denen die einzelnen Altersstufen zu prüfen wären, in langwieriger empirischer Arbeit ausgesucht werden. Sie wären nämlich für jede Altersstufe so zusammenzustellen, daß der durchschnittliche Seltenheitswert aller vorgelegten Tests 25% beträgt. Denn dies entspräche dem bekannten Prinzip, daß die Lösungshäufigkeit der Tests durchschnittlich 75% betragen solle. Somit ist das, was

Schultze durch sein Maß glaubte vermeiden zu können: die besondere Anpassung der Tests an die bestimmten Altersstufen, nicht nur ebenso erforderlich wie bei der Bestimmung des IA., sondern bereitet noch höhere Schwierigkeiten, da für jede Alterstufe nach Schultzes Vorschlag 25 Tests zu einer geeigneten Serie vereint werden müßten. Außerdem würde es hier nicht genügen, für je einen Jahrgang eine solche Serie aufzustellen; vielmehr müßte man auch für Zwischenstufen Gleiches tun, da ja die Leistung eines Kindes nur durch den Vergleich mit wirklich Gleichaltrigen ihren Seltenheitswert erhält.

Ein weiterer und wohl der hauptsächliche Nachteil der IZ. besteht darin, daß sie für die außerhalb der Normalität liegenden Intelligenzen keine Abstufungen mehr ergibt. Die übernormalen Kinder erhalten durchweg die IZ. 1, sobald sie alle Tests, die ihrer Altersstufe vorgelegt werden, gelöst haben; etwaige Lösungen von Tests noch höherer Altersstufen können ihnen nicht angerechnet werden. Und die Schwachsinnigen erhalten alle die IZ. 0, sobald sie keinen der für ihr Alter vorgesehenen Tests lösen; die tieferen Staffeln, deren Tests sie etwa noch zu lösen imstande sind, können ihnen nicht angerechnet werden. Aber selbst, wenn sie eine ganz geringe Anzahl der für ihr Alter bestimmten Tests lösen, so ist es doch ein großer Mißstand, daß ihre IZ. aus ganz wenigen Werten abgeleitet werden soll, während bei den befähigteren Kindern ein Durchschnittswert aus viel mehr Komponenten gebildet wird.

Somit scheint die IZ. nicht geeignet zu sein, die von Schultze in dies Maß gesetzten Hoffnungen zu erfüllen und an die Stelle des I.-Quotienten zu treten, dem die oben angegebenen Mängel nicht anhaften. Dagegen scheint mir das Prinzip des Seltenheitswertes recht beachtenswert; es dürfte für manche enger begrenzte Aufgaben gute Dienste leisten.

Neuerdings hat Lipmann (*198*) dieses Prinzip weiter ausgebaut, indem er es nicht nur, wie Schultze, auf Alternativtests beschränkt. In abstufbaren Tests wird dann eben jeder Bestandteil besonders dieser Berechnung unterzogen und daraus ein Resultantenwert gewonnen. Ist L die Lösung, die ein solcher Bestandteil eines Tests gefunden hat, so ist deren „Schwierigkeitsindex" (SI.) „diejenige Zahl, die angibt, wieviel Prozent der Vpn. eine Lösung gefunden hat, die schlechter ist als die Lösung L". In einem Lückentext z. B. wird für jede einzelne Lücke gezählt, wieviel Prüflinge sie ganz richtig ausgefüllt haben. Sind dies für eine Lücke z. B. 30%, so ist der Schwierigkeitsindex der richtigen Ergänzung = 70%. Haben 15% andere die Lücke mit einem anderen noch zulässigen, aber nicht so hoch zu bewertenden Wort ausgefüllt, so ist für diese Lösung SI. = 55% (denn 55% haben eine noch schlechtere Leistung aufzuweisen); haben weitere 10% ein nur notdürftig passendes Wort gefunden, so ist dessen SI. = 45% usw.

Hier ergibt sich schon eine bedeutende methodische Schwierigkeit: liegen für irgend eine Einzelaufgabe verschieden lautende Lösungen vor, so müssen diese in eine Güte-Rangordnung gebracht werden, damit festgestellt werden kann, wieviele Lösungen schlechter sind als irgend eine bestimmte Lösung. Diese Rangierung wird stets in gewissem Maße willkürlich sein und zuweilen wird diese Willkür einen Grad erreichen, der die wissenschaftliche Brauchbarkeit in Frage stellt. L. gibt selbst zu, daß die Einreihung der Null-Lösungen (wo überhaupt keine Antwort erfolgte) unüberwindliche Schwierigkeiten mache, da man nicht wisse, ob die Unterlassung einer Angabe höher oder geringer zu bewerten sei als eine fehlerhafte Angabe. Er läßt die Null-Lösungen daher ganz aus der Berechnung heraus. Aber auch bei den positiven Lösungen ist die Bewertung oft stark in das Belieben des Prüfers gestellt. Es gibt Tests, wo fast jeder Prüfling eine anderslautende Lösung findet, z. B. wenn er aus einer gegebenen Fabel die Moral ziehen soll. L. versucht auch in solchen Fällen eine Güte-Rangordnung der Lösungen herzustellen, um daraus für jede den SI. zu berechnen; aber nach den von ihm selbst gegebenen Listen der Lösungen und Bezifferungen erscheint es mir fraglich, ob nicht jeder andere Experimentator eine andere Rangordnung vorgenommen hätte.

Die Schwierigkeitsindizes bilden nun nach Lipmann den Ausgangspunkt weiterer Berechnungen. Bei den Tests, die aus vielen Einzellösungen bestehen (z. B. Lückenergänzungen), wird für jeden Prüfling aus sämtlichen Schwierigkeitsindizes der von ihm gebrachten Lösungen ein „Funktionsindex" für diesen Test gewonnen. Dies geschieht, indem die Einzelindizes addiert und die Summen mit Prozentzensuren versehen werden. Die Funktionsindizes sämtlicher geprüfter Tests werden dann nochmals auf gleiche Weise für jeden Prüfling zu Resultantenwerten vereinigt, die L. als „I.-Zensuren" bezeichnet.

VI. Kapitel.

Prüfung der einzelnen Intelligenzfunktionen.

Die folgende Besprechung der einzelnen prüfbaren I.-Gebiete schließt sich eng an die soeben erschienene „Methodensammlung zur Intelligenzprüfung von Kindern und Jugendlichen" (*8*, III) an. Unter Hinweis auf diese Sammlung kann hier die Aufzählung und Beschreibung der Tests selbst erspart werden; dagegen soll in Kürze die dort fehlende psychologische Beurteilung der Hauptgruppen erfolgen. Hierbei halten wir uns an die tatsächlich vorgeschlagenen und angewandten Tests, obwohl wir uns dessen bewußt sind, daß deren Liste höchst ungleichwertig ist und daß einige Gebiete unverhältnismäßig stark bedacht, andere bisher recht vernachlässigt worden sind.

Vorauszuschicken ist, daß die Einteilung und Benennung der Einzelgebiete ebenso wie die Einordnung der Tests in sie nicht ohne Willkür möglich ist. Es gibt keinen Test, der nur die „Merkfähigkeit" oder die „Kritikfähigkeit" usw. prüfte, ohne daß irgend eine andere geistige Funktion dabei beteiligt wäre; bei manchen Tests kann man geradezu zweifelhaft sein, welchem Gebiet sie mit mehr Recht zuzuzählen seien. Die Unterstellung bestimmter Tests unter bestimmte Stichworte soll also nur bedeuten, daß die Tests vorwiegend zur Feststellung eben jener Einzelfunktionen geeignet erscheinen.

Die Tests lassen sich nach ihrer Beziehung zur I. in mittelbare und unmittelbare scheiden.

Bei den mittelbaren ist die Aufgabe selbst nicht eigentlich intelligenz-artiger Natur; sie geht auf die Erzielung einer Gedächtnis-, einer Aufmerksamkeits-, einer Reaktionsleistung, auf die Erzeugung von räumlichen Vorstellungen und Darstellungen usw. I.-Akte sind lediglich als Mittel zur Erreichung dieser Ziele beteiligt; und nur insofern haben die Tests für uns Bedeutung.

Die Entwicklung der Testforschung brachte es mit sich, daß man mit diesen mittelbaren Methoden begann. Man hat eben zuerst all jene Verfahrungsweisen, die dem psychologischen Experimentator von anderen Problemen her vertraut waren, auf unsere Fragestellung übertragen, indem man einfach untersuchte, ob bei diesen Messungen der Reaktionszeit, bei Bestimmungen der Unterschiedsempfindlichkeit, bei Assoziations-, Lern- und Merkversuchen, bei Geschicklichkeitsmessungen usw. intelligentere Personen bessere Ergebnisse lieferten als weniger intelligente. Wollen wir diese Methoden noch weiter anwenden, so müssen wir untersuchen, an welcher Stelle der Leistung etwa ein Intelligenzanteil stecke, welcher Art er sei, und ob er neben den anderen beteiligten Funktionen genügend zur Geltung komme, um die Benutzung des Tests als I.-Test zu rechtfertigen.

Allmählich erkannte man die Unzulänglichkeit solcher mittelbaren Prüfungen und erfand Tests, die unmittelbar intelligente Zielhandlungen vom Prüfling verlangten. Ebbinghaus' Aufstellung des Kombinationstests (Textlückenergänzung) (*60*) war hier bahnbrechend. Andere kombinatorische Aufgaben folgten, und immer mehr eigentliche Denktätigkeiten wurden in den Bereich der Testprüfung gezogen: das Verstehen, Erfassen des Wesentlichen, Vergleichen, Definieren, Urteilen, Kritisieren, Schließen usw.

Die Entwicklung der letzten Jahre hat gerade diese Einbeziehung der höheren Denktätigkeiten begünstigt, da das neue Problem der Begabtenauslese mit den früheren Tests, die zu leicht waren und die I. nicht unmittelbar genug erfaßten, nicht auskam.

Bereits werden einige Richtlinien für die weitere Entwicklung sichtbar: Man erkennt jetzt, daß wir bisher zu einseitig sprachlich-logische Tests angewandt hatten, und bemüht sich um die Ausbildung stummer Tests, die die praktische Handlungsfähigkeit (das Ordnen, das Ausführen von Aufträgen, die konstruktive Phantasie, die Findigkeit usw.) prüfen sollen. Ferner drängt sich immer mehr die Notwendigkeit auf, neben der allgemeinen Begabung auch bestimmte Sonderbegabungen zu prüfen, da die Feststellung des sprachlichen, mathematischen, technischen, zeichnerischen usw. Talentes sowohl für die Zuweisung zu bestimmten Schulen wie für die Aufnahme in Lehrlingswerkstätten und für die Berufsberatung wichtig ist; derartige Spezialbegabungsprüfungen stehen mit den IP. in engem Zusammenhang.

Auch diese in die Zukunft weisenden methodischen Richtungen werden in der folgenden Betrachtung behandelt werden, soweit sie schon zu greifbaren Methodenvorschlägen geführt haben.

Reihenfolge und Zählung der Einzelgebiete entspricht der Einteilung der „Methodensammlung“[1]).

1. Kenntnisprüfungen (MS.[1]) 8—13).

Diejenige Verfahrungsweise, die am meisten an die sonst üblichen nicht-psychologischen Prüfungen, insbesondere Schulexamina erinnert, ist die Feststellung von Kenntnissen. Die Einsicht, daß Kenntnisse nicht mit Fähigkeiten verwechselt werden dürfen, ist die Vorbedingung für eine fruchtbare Entwicklung der psychologischen Prüfungsmethoden gewesen. Ziemlich früh führte diese Einsicht dazu, daß man eigentliches Schulwissen aus den Prüfungszielen ausschaltete; denn dessen Erwerb ist ja in sehr hohem Maße und vor allem in unkontrollierbarer Weise von äußeren Bedingungen (Art der Beschulung, des Lehrplans und der Unterrichtsmethoden, Schulwechsel, häuslicher Hilfe usw.) abhängig, so daß ein eindeutiger Schluß auf Fähigkeiten nicht möglich ist. Natürlich ist es nicht zu vermeiden, daß man gewisse in der Schule

[1]) Im folgenden ist der Hinweis auf die Methodensammlung durch die Abkürzung MS. bezeichnet. Die Ziffern dahinter bezeichnen die Seitenzahl.

erworbene Kenntnisse und Fertigkeiten, z. B. Lesen und Schreiben, Kenntnis bestimmter Ausdrücke usw., als Mittel zur Lösung eines Tests voraussetzt; hier muß als Norm dienen, daß nur solche Schulkenntnisse vorausgesetzt werden dürfen, die für die betreffende Gruppe oder Altersstufe als selbstverständlich und nicht von der Begabung abhängig gelten können[1]).

Etwas anders steht es um die Kenntnisse des Alltagslebens, die unabhängig von der Schule dem Kinde gleichsam anfliegen. Binet und seine Nachfolger haben solche Fragen ihrem Prüfungssystem eingeordnet; so wird bei 3—6jährigen nach dem eigenen Namen, dem Namen bekannter Gebrauchsgegenstände und Körperteile, nach der Bezeichnung der Hauptfarben usw. gefragt; etwas ältere Kinder sollen die Wochentage und Monate, die Tageszeiten (Unterschied von Vormittag und Nachmittag), die geläufigen Münzen und die einfachsten Hantierungen mit ihnen, die Zahl der Finger usw. kennen.

Daß solche spontan erworbenen Kenntnisse schon mehr mit der I. zu tun haben, als das durch die Schule aufgezwungene Wissen, ist wohl sicher. Das Kind nimmt ja nicht rein passiv die sich ihm darbietenden Eindrücke und Namen auf, sondern kommt ihnen entgegen; die geistige Regsamkeit, das intellektuelle Interesse, die leichtere oder trägere Auffassung wirken in hohem Grade dabei mit, daß diese Kenntnisse von einem Kinde früher, von dem anderen später erworben werden.

Aber daneben spielt doch auch das Milieu eine sehr bedeutende Rolle, insbesondere die im Hause gebotenen Anregungen und Gelegenheiten, um Dinge und Bezeichnungen und ihre richtige Anwendung kennenzulernen. So ist der Sprachschatz von Kindern aus gebildetem Hause meist sehr viel größer als der gleichaltriger Kinder aus einfachem Stande, ohne daß diese Abweichung entscheidend mit der I. zusammenhinge; umgekehrt wissen arme Kinder, die den Eltern Besorgungen abnehmen und im Geschäft

[1]) Die geringe Verläßlichkeit von Schulkenntnisaufnahmen für die IP. Erwachsener ist im Jahre 1905 von Rodenwaldt (*96, 97*) mit schlagender Deutlichkeit nachgewiesen worden. Die Psychiater hatten früher zur Feststellung geistiger Minderwertigkeit ganz vorwiegend Kenntnisprüfungen verwandt, einfachste geographische, historische, heimatkundliche, rechnerische Fragen; R. prüfte nun mit diesen Fragen eine große Anzahl geistesgesunder Rekruten und stellte eine ganz erschreckende Unkenntnis fast aller Prüflinge fest. Ein Schluß von diesen mangelnden Kenntnissen auf die geistige Minderwertigkeit der Prüflinge wäre also unberechtigt gewesen.

helfen, oft viel besser, Münzen zu benennen und mit ihnen umzugehen, als Kinder aus wohlhabendem Hause, die von allen Geldingen ferngehalten werden.

So sind denn auch Kenntnisse des Alltagslebens nur mit großer Vorsicht und Zurückhaltung im Rahmen von IP. zu verwerten.

2. Prüfung der Merkfähigkeit (MS. 13—40).

Die Fähigkeit, einmal dargebotene Inhalte unmittelbar nach der Darbietung richtig zu reproduzieren oder wiederzuerkennen, ist im wesentlichen von drei Funktionen abhängig. Erstens ist es von Bedeutung, mit welcher Leichtigkeit und Deutlichkeit der sinnliche Eindruck des Wortlautes oder der optischen Form über den Augenblick der Darbietung hinaus im Bewußtsein haftet (Perseveration; unmittelbares mechanisches Gedächtnis). Zweitens muß die Aufmerksamkeit genügenden Umfang haben, um die vorgelegten Inhalte nach Zahl und Anordnung zu erfassen, und genügende Ausdauer, um den erfaßten Stoff bis zur Reproduktion festzuhalten. Endlich spielt der Umstand mit, ob das Dargebotene über den bloß sinnlichen Eindruck hinaus gedanklich erfaßt und dann auch gedanklich reproduziert wird (Akt des „Verstehens"; unmittelbares judiziöses Gedächtnis).

Von diesen drei Funktionen ist die erste für die I.-Frage am wenigsten wichtig; die rein mechanische Perseveration und die Fähigkeit des Nachplapperns sind auch bei geistig sehr tief stehenden Individuen oft recht ausgebildet. Deshalb ist das Arbeiten mit sinnlosen Silben, so wertvoll es für Probleme der eigentlichen Gedächtnispsychologie sein mag, für I.-Untersuchungen von geringer Bedeutung (MS. 13—14). Es kann höchstens sekundär dort in Betracht kommen, wo man gewisse Seiten der Sprachbegabung (Fähigkeit des Vokabelmerkens und -lernens) untersuchen will.

Der zweite Faktor, die determinierende Tendenz der Aufmerksamkeit, scheint in besonderem Maße eine Funktion des geistigen Altersfortschritts zu sein; denn der Umfang des behaltbaren Stoffes wächst aufs deutlichste mit steigendem Alter. Es gibt wohl kaum eine geistige Leistung, die so deutlich vom Alter abhängig wäre, wie das Merken einer Vorstellungsreihe bestimmten Umfanges; das ist für Ziffern von Bobertag und anderen (MS. 14—15), für Sätze insbesondere von Gaßmann und Schmidt nachgewiesen worden (MS. 24—35).

Bobertag (*51*) fand, daß im allgemeinen 5jährige Kinder vier einmal vorgesagte Ziffern, 7jährige Kinder fünf Ziffern, 10jährige sechs Ziffern richtig wiederholen können.

Aus Gaßmanns und Schmidts Tabellen (*62*) lassen sich für die ziemlich schweren von ihnen verwandten Sätze (Satzverbindungen und Satzgefüge) folgende Eichungswerte aufstellen:

Lebensalter	7.	8.	9.	10.	11.	12. Jahr
richtig zu wiederholende Satzlänge	14	16	19	20	21	22 Silben.

Der behaltbare Umfang steigt also zuerst schnell, dann langsamer.

Die eindeutige Beziehung des Merkumfangs zum Alter wird noch verstärkt durch den anderen Befund Gaßmanns und Schmidts, daß gleichaltrige Kinder verschiedener I. hierin geringe Unterschiede zeigen. Die Berücksichtigung des Merkumfanges bei IP. ist also nur gerechtfertigt zur Feststellung von Entwicklungsverzögerungen, also eigentlicher „Zurückgebliebenheit"; hierfür leistet sie aber treffliche Dienste.

Viel unmittelbarer mit der IP. hat es der dritte Faktor zu tun, die Beteiligung des Verstehens am Merkprozeß. Jeder Merkstoff, mit Ausnahme von sinnlosen Silben und sinnlosen geometrischen Figuren (MS. 16—21), hat außer seinem sinnlich wahrnehmbaren Material noch einen gedanklichen Gehalt: Bedeutungen der einzelnen Wörter bzw. bildlichen Darstellungen, sinnvolle Beziehungen der einzelnen Elemente untereinander, Gesamtbedeutung des Sinnganzen. Beim Intelligenten geht schon das Erfassen des dargebotenen Stoffes so vor sich, daß jene gedanklichen Inhalte sich teils von selbst einstellen, teils, wo sie zunächst fehlen, ausdrücklich gesucht werden. Seine geistige Regsamkeit vermag sich eben nicht zu beruhigen beim bloß mechanischen Wort- oder Figurenstoff. Und wenn er dann reproduziert, so stellt sich neben den mechanisch perseverierenden Lauten und Linien, ja zuweilen vor diesen und unabhängig von diesen, auch jener gedankliche Inhalt wieder ein und erleichtert die Wiedergabe. Verstandenes behält sich besser als Unverstandenes; und darum springt beim judiziösen Behalten sofort der Unterschied der Intelligenzen hervor, der beim rein mechanischen Behalten nicht erkennbar ist.

Die Prüfungsmethoden lassen diesen Anteil des Gedanklichen in verschiedener Weise hervortreten.

Gaßmann und Schmidt fanden, daß die rein silbenmäßige Länge der Sätze bei Intelligenten viel weniger für die Merkfähigkeit bestimmend ist als bei Unbegabten. Die letzteren müssen hier eben ihre Aufmerksamkeit auf die mechanische Bewältigung einer

bestimmten Silbenzahl einstellen; jene dagegen merken sich die Gedankeneinheiten und Zusammenhänge und rekonstruieren von hier aus wieder den Wortlaut, dessen Silbenzahl daher für die Behaltbarkeit nicht so entscheidend sein kann.

Meumann (*85*) (MS. 21) fand, daß bei Darbietung zusammenhangloser Wortreihen der Unterschied der Intelligenzen besonders dort hervorsprang, wo abstrakte Worte dargeboten wurden. Die weniger Intelligenten brachten für diese schweren Worte noch nicht das rechte Verständnis auf und konnten sie daher auch schwerer behalten.

Noch deutlicher wird der Leistungsunterschied, wenn nicht Einzelworte, sondern Zweier- oder Dreiergruppen von Worten dargeboten werden, deren jede von einem logischen Verknüpfungsprinzip beherrscht wird; z. B.: Tauwetter — Hochwasser; Fleiß — Lob; Arbeit — Verdienst — Wohlstand usw. Nach Darbietung einer Reihe solcher Gruppen wird das erste Wort jeder Gruppe wiederholt, der Prüfling muß nun das oder die anderen Worte reproduzieren. Hier wird also der Beziehungsgedanke nicht mit dargeboten, sondern muß gleich beim Hören vom Prüfling gefunden werden, damit er ihn als Reproduktionshilfe verwerten kann. Geschieht dies nicht, so ist er auf rein mechanische Reproduktion der Wortreihe angewiesen; diese aber vermag er bei einer größeren Zahl von Wortgruppen nicht zu leisten.

Dies von Ranschburg (*92*), Ries (*95*) und Weigl (*122*) ausgebildete Verfahren (MS. 21—24) ergab eine sehr hohe Korrelation mit der nach Lehrerschätzung hergestellten I.-Reihe der Prüflinge (vgl. Kap. X, 3) und gilt heute als eine der zuverlässigsten Prüfungsmethoden.

Von besonderem psychologischen Interesse sind diejenigen Merkleistungen, bei denen sich der gedankliche Gehalt von dem sinnlichen Merkstoff mehr oder weniger unabhängig macht. Hier haben wir nämlich den Beweis für die von der modernen Denkpsychologie behaupteten Existenz selbständiger, von der Anschauung unabhängiger „Gedanken“[1]).

Bei IP. tritt diese Verselbständigung des Gedanklichen dort hervor, wo das rein mechanische Merken wegen zu großer Länge

[1]) Zur Veranschaulichung dieses grundsätzlich wichtigen Tatbestandes diene folgendes fingierte Beispiel. Y. sagt einem Boten: „Bestellen Sie bitte Herrn X., ich fühle mich heute leidend und kann deshalb der Sitzung nicht beiwohnen.“ Und der Bote richtet aus: „Herr Y. kann nicht in die Versammlung kommen,

des Merkstoffes versagt. Der Prüfling muß — im Sinne des wörtlichen Merkens — Fehler machen; entscheidend ist nun, ob es bei den so entstehenden Lücken und Sinnlosigkeiten verbleibt oder ob das Denken einspringt, um in gleicher oder ähnlicher Bedeutung — jedenfalls in sinnvoller Weise — den Satz zu rekonstruieren. Gaßmann und Schmidt stellten fest, wie häufig bei nichtgelungener wörtlicher Wiedergabe von Sätzen sinnlose Fehler und sinnvolle Fehler begangen wurden, und fanden, daß das Verhältnis beider Fehlerarten in starker Korrelation sowohl zu Alter wie zu Intelligenz der Prüflinge stand.

Der Prozentsatz der sinnlosen Fehler unter allen Fehlern betrug bei den 7jährigen fünfmal soviel wie bei den 13jährigen, bei den Schwachbegabten aller Altersstufen im Durchschnitt doppelt soviel wie bei den Begabten.

Der hier vorliegende Gesichtspunkt schien uns so wertvoll, daß wir ihn bei der Hamburger Begabungsprüfung (*218*, 10) in den Mittelpunkt einer besonderen Versuchsanordnung stellten. Die zum Merken aufgegebenen Sätze wurden so lang gewählt, daß ein wörtliches Behalten ausgeschlossen schien; den Prüflingen wurde aufgegeben, die Sätze niederzuschreiben, „so gut sie sie behalten hatten" (MS. 36). Es ließ sich ohne weiteres feststellen, ob der Versuch zu sinnvollem Ersatz des nicht wörtlich Behaltenen gemacht wurde oder nicht. Die Ergebnisse zeigten wiederum hohe Korrelation mit anderweitigen Urteilen über die I. der Kinder.

Am meisten entfernen sich die Aufgaben dieser Gruppe von bloßen mechanischen Gedächtnisaufgaben dort, wo Gesichtspunkte aus einem größeren Zusammenhange gemerkt werden sollen (MS. 37—40). Wenn eine gehörte oder gelesene Geschichte wiederzugeben ist, so wird von vornherein nicht erwartet, daß der ganze Inhalt behalten werden kann. Vielmehr nimmt schon die Auffassung beim Aufnehmen, und dann wiederum die Erinnerung beim Wiedergeben eine doppelte intellektuelle Bearbeitung vor: aus dem Gesamtstoff wird eine Auslese getroffen, und zwischen diesen Elementen wird ein Zusammenhang hergestellt. Die Ökonomie des Gedächtnisses fordert, daß bei allem Wahrgenommenen sehr viel weniger gemerkt und behalten wird, als dargeboten wurde;

---

weil er nicht ganz wohl ist." Dann hat sich der Bote den Auftrag seinem Gehalt nach vollständig richtig gemerkt, aber kein einziges der Worte, die ihm gesagt worden waren, sprachlich reproduziert. Ein mechanisches Merken des Anschauungsstoffes fehlt also vollkommen; die Gedanken sind rein als solche gemerkt worden.

von der I. aber hängt es ab, ob das Behaltene nur ein Mosaik zufällig gemerkter Bruchstücke oder eine Verdichtung des Gegebenen auf die wesentlichen Bestandteile und auf die entscheidende Beziehung zwischen diesen darstellt. Hier leiten die Merkfähigkeitsprüfungen unvermerkt über in andere, welche unmittelbar auf intellektuelle Zielsetzung gehen[1]).

3. Prüfung des räumlichen Vorstellens (MS. 41—60).

Die Prüfung des räumlichen Vorstellens, wozu auch aus dem vorigen Abschnitt die Merkfähigkeitsprüfungen für Formen und Bilder zu rechnen sind (MS. 16—21), haben zu ihrem ursprünglichen Ziel die Feststellung einer Sonderbegabung, der Visualität, das ist die Fähigkeit, räumliche Gestalten, ohne daß sie äußerlich gegeben wären, vorzustellen, oder gegebene Formen in der Vorstellung zu zerlegen, umzuwandeln, aus Teilen zusammenzusetzen und endlich die Ergebnisse dieser optischen Vorstellungsarbeit graphisch darzustellen — diese Fähigkeit findet sich oft ziemlich isoliert, ohne notwendige Beziehung zu einem bestimmten I.-Grad oder zu anderen Sonderfähigkeiten[2]).

Eine Prüfung der Visualität als solcher ist für eine Reihe von Berufen von Bedeutung, so für alle, die es mit räumlichen Gestaltungsaufgaben, Zeichnungen, technischen Konstruktionen zu tun haben. Ein großer Teil der in der „Methodensammlung" geschilderten Methoden stammt auch aus neueren Eignungsprüfungen für technische Zwecke; sie sollen der Auslese von Lehrlingen für die Werkschulen technischer Betriebe dienen (MS. 47—49 und 53—56).

Aber die meisten der einschlägigen Aufgaben sind auch so gefaßt, daß sie zugleich die I. in Bewegung setzen. Ja man kann sagen, daß gerade dort, wo die besondere visuelle Begabung nicht stark ausgebildet ist, bewußte Überlegung, Konstruieren in Gedanken sie teilweise zu ersetzen vermag. Ob freilich die endliche Lösung einer solchen Aufgabe wesentlich durch die natürliche Anschauungsbegabung oder durch die Fähigkeit, sich intellektuell

---

[1]) Vgl. zum letzten insbes. Abschnitt 18 dieses Kapitels: Finden des Wesentlichen (S. 116).

[2]) Zuweilen kommt sogar eine starke visuelle Fähigkeit bei geistiger Minderwertigkeit vor. Der „Katzenraphael" war ein imbeziller, in der Darstellung von Katzen außerordentlich gewandter Maler.

zu helfen, zustande gekommen ist, läßt sich aus den fertigen Ergebnissen meist nicht mehr entnehmen; dagegen vermag die Beobachtung des arbeitenden Kindes selbst darüber Klarheit zu verschaffen. In dieser Hinsicht ist die Arbeit von Köhn (*75*) vorbildlich, der geometrische Figuren aus 2, 3 oder 4 Bruchstücken zusammensetzen ließ, nachdem vorher eine vollständige Figur als Muster gezeigt worden war (MS. 42—45).

K. schildert uns die typische Verschiedenheit des intelligenten und des unintelligenten Kindes bei dieser Aufgabe. Schon bei der Auffassung der Vorlage tritt der Begabungsunterschied hervor, indem das begabte Kind von selbst das Wesentliche sich zu vergegenwärtigen sucht, die Ecken abzählt und hierbei ohne äußere Anregung auf die nötige Hilfe (z. B. Festhalten der Ausgangsecke) kommt, eine zutreffende Bezeichnung findet, das Muster auf die dann zu lösende Kombinationsaufgabe bezieht. Auch beim Zusammensetzen selbst zeichnet sich das intelligente Kind durch das Vorherrschen von Eigenimpulsen, durch dauernd wirksame determinierende Tendenz, Auslese des Passenden, Ausnützen günstiger Chancen aus; der ganze Kombinationsvorgang ist ein zusammenhängender zielstrebiger Prozeß, während er beim unintelligenten in eine Reihe isolierter Akte des tastenden Herumprobierens zerfällt.

Man erkennt schon hieraus, daß bei diesen und verwandten Tests die Beachtung des qualitativen Verhaltens während der Arbeit fast wichtiger für die Schätzung der Intelligenz ist als der schließliche Erfolg, ob die Aufgabe richtig oder falsch gelöst worden ist. Solche Tests sind natürlich besonders wertvoll; sie sind in der Hand eines psychologisch geschulten und gut beobachtenden Prüfers ein Hilfsmittel, das über das nackte Zahlenmaterial der „+“ und „—“ hinaus ein Bild von der I.-Beschaffenheit des Kindes gewährt.

Statistisch ergibt sich aus den Versuchen K.s, die freilich nur an einer kleinen Anzahl von Kindern aus wenigen Altersstufen angestellt wurden, eine enge Korrelation mit Allgemeinbegabung und Alter. Von 35 Figuren, welche den 8jährigen vorgelegt wurden, wurden glatt gelöst: von den Gutbegabten 26, den Mittelbegabten 17, den Schwachbegabten 8. Bei den 6jährigen lauten die entsprechenden Zahlen: 20, 10, 1.

Auch bei den anderen Aufgaben aus dem Gebiet des räumlichen Vorstellens ist eine solche Beteiligung intellektueller Tätigkeiten feststellbar. So gibt Rossolimo (*99*) die Aufgabe (MS. 45),

einige vorgelegte Figuren aus einer Anzahl von Quadraten und Dreiecken zusammenzusetzen; wer sich nun etwa bemüht, eine Raute mit schiefen Winkeln lediglich durch Zusammensetzung aus kleinen Quadraten zustande zu bringen, hat nicht nur geringe Anschauungsfähigkeit, sondern ihm fehlt Überlegung und Umsicht. Andere Proben erfordern von dem weniger visuell Begabten einen bedeutenden Aufwand an geistiger Energie: es gilt die spröden, leicht entgleitenden Formvorstellungen festzuhalten und allmählich zu vervollständigen oder umzuwandeln; hierbei muß die determinierende Tendenz der Aufgabe dauernd gegen Erschlaffung und Ablenkung aufrechterhalten bleiben. (Beispiele: Uhrzeigervertauschung (MS. 52), Entfaltungsversuch (MS. 49—50)[1]). Dann wieder müssen abstrakte Überlegungen zu Hilfe genommen werden: so bei den Fragen nach der Zahl der Diagonalen des Würfels und des Fünfecks (MS. 53) oder beim Nachzeichnen von Figuren (MS. 59—60), wo man durch die Begriffe symmetrisch und asymmetrisch, rechtwinklig und schiefwinklig, rechts länger, links kürzer usw. das Gedächtnis unterstützen kann. Eine andere Gruppe von Aufgaben setzt die allgemeine Fähigkeit des Vergleichens und Unterscheidens, des Ordnens und Zuordnens voraus, so das Einpassen der Klötze in die Vertiefungen des Formenbretts (MS. 56—57), das Ordnen von Werkstücken nach Größe, Dicke usw. (MS. 56), der Vergleich eines durchlochten Blechstückes mit zehn untereinander ähnlichen Zeichnungen (MS. 17—18) usw.

Eine von mir angegebene schwierige Visualisationsaufgabe wird zur Zeit im Hamburger Laboratorium durchgeprüft; zur Ergänzung der Materialien der „Methodensammlung" sei sie hier kurz geschildert. Aufgabe: Der Prüfling stelle sich einen Würfel vor, dem eine Kugel eingeschrieben ist (die also den Würfel an den Seitenmitten berührt) und eine Kugel umschrieben ist (die ihn an den Ecken berührt). Nun wird das ganze System durch einen Schnitt halbiert. Wie sieht die Schnittfläche aus? (Zeichnen!) Die Lage des Schnittes kann nun mit steigender Schwierigkeit vorgeschrieben werden. Nämlich:

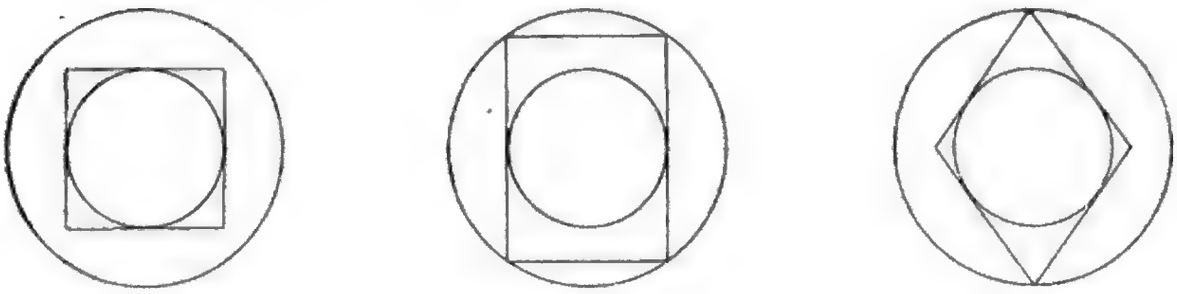
a) parallel zu zwei Würfelseiten,
b) schräg von der rechts oberen zur links unteren Kante, so daß diese Kanten selbst in der Schnittfläche liegen,

---

[1]) Auch bei dem Abelson-Test (*46*) (Erkennen der Lagebeziehungen sich kreuzender geometrischer Figuren, MS. 58—59) scheint es sich vor allem um eine hohe Anforderung an die Aufmerksamkeit zu handeln, da hier eine Reihe von Teilzielen zunächst einzeln klar gemacht, dann miteinander kombiniert werden müssen; doch ist die psychische Struktur dieses Tests noch wenig untersucht.

o) schief von einer Ecke entlang der Raumdiagonale zur anderen Ecke, so daß die Raumdiagonale die Achse der Schnittfläche bildet.

Die drei Schnittfiguren haben folgendes Aussehen:



4. Beurteilung technisch-physikalischer Verhältnisse (MS. 60—65).

Vorstellen von Verschiebungen im Raume.

Die hierher gehörigen Tests dienen bisher ausschließlich einer Spezialaufgabe, nämlich die Eignung für technische Berufe festzustellen; aber es ist vielleicht empfehlenswert, auch den Serien für die allgemeine IP. gelegentlich einen solchen Test einzureihen, um neben der sprachlich-theoretischen auch die praktisch-technische Seite der I. hervortreten zu lassen.

Dem Prüfling wird eine ganz einfache technische Vorrichtung gezeigt mit der Anweisung, Bewegungsverlauf und -bedeutung anzugeben. Wichtig ist, daß das Modell keine Spezialkenntnisse und spezielle Übung auf technischem Gebiet voraussetzen darf; man muß bei genügender Überlegung aus der Konstruktion des Modells selbst ablesen können, welche Verrichtung es vollziehen soll. Geeignet sind hierfür Bewegungsübertragungen durch Zahnräder, Treibriemen, Scharniere und Hebel.

Psychologisch kommen zunächst alle im vorigen Abschnitt geschilderten Faktoren in Betracht, dazu aber als wichtiges neues Moment die Fähigkeit, Bewegungen und Bewegungsverknüpfungen vorzustellen. Es ist wohl denkbar, daß es Menschen mit starker Visualität gibt, die aber nur ruhende Raumgebilde vorstellen können (z. B. Forscher auf dem Gebiet der Geometrie). Im Gegensatz hierzu ist beim Techniker gerade die Umwandlung der ruhenden simultanen Raumvorstellung in eine Abfolge von Raumvorgängen, die auseinander hervorgehen, charakteristisch. Er sieht in einem ruhenden Zahnrad mit eingreifender Zahnstange nichts anderes als den Ausgangspunkt zweier notwendig aufeinander

bezogener Bewegungen; der funktionale Zusammenhang ist für ihn geradezu eine Anschauungstatsache geworden.

Die Fähigkeit, bewegte Raumgebilde anschaulich vorzustellen, kann auch ganz unabhängig von irgend welchen technischen Problemen geprüft werden. Einige von mir vorgeschlagene Tests dieser Art („Verschiebungstests") werden zur Zeit im Hamburger Laboratorium durchgeprüft; da sie in der „Methodensammlung" noch nicht enthalten sind, seien sie hier in Kürze geschildert. Der Prüfling soll sich Figuren vorstellen, die aus starren, durch Kugelgelenke verbundenen Stäben gebildet sind. Die Stäbe können an den Verbindungsstellen gegeneinander beliebig gedreht, aber in ihrer Länge nicht verändert werden. Nun soll der Prüfling aus der Vorstellung heraus für verschiedene zwei- und dreidimensionale Figuren solcher Art angeben, ob und wie sie durch Verschiebung der Stäbe in ihrer Form geändert werden können.

1. Dreieck. (Richtige Antwort: kann nicht verschoben werden.)

2. Viereck. (Kann verschoben werden, und zwar bis zu einem Dreieck, wenn zwei benachbarte Seiten nicht die gleiche Längensumme haben wie die beiden anderen Seiten, sonst bis zu einer Geraden.)

3. Pyramide. (Kann nicht verschoben werden.)

4. Würfel. (Hier sind Vorstellungen sehr verschiedener Schwierigkeit zu vollziehen. Ein Würfel mit der Kantenlänge a kann zunächst umgeklappt werden zu einem Rechteck mit den Seiten $2a$ und a; dies kann dann wieder bis zu einer Geraden von der Länge $3a$ zusammengedrückt werden. Der Würfel kann aber auch direkt längs einer Raumdiagonale auseinandergezogen werden zur Länge $3a$ oder zusammengedrückt werden zu einem Sechseck. Diese letzte Verschiebung ist am schwersten vorstellbar.)

5. Ausführung technischer Aufgaben (MS. 65—70).

Praktische Intelligenz.

Waren im vorangegangenen Fall bestimmte technische Mittel gegeben, aus denen der Erfolg gedanklich abgeleitet werden sollte, so wird nun das Ziel aufgegeben, und die geeigneten Mittel und Wege sollen gefunden und — was noch mehr ist — angewandt werden, so daß das Ziel erreicht wird. Auch hier dürfen wiederum keinerlei Kenntnisse und Fertigkeiten spezieller Art vorausgesetzt werden.

Untersucht wird, ob ein Zielgedanke, verbunden mit der Anschauung der etwa zur Verfügung gestellten einfachen Materialien, die richtige determinierende Tendenz auf die Überlegung und das Handeln ausübt. Es muß wiederum eine Auslese unter den verschiedenen Möglichkeiten des Tuns getroffen werden; es müssen diese Möglichkeiten nach ihrer Brauchbarkeit beurteilt werden; es müssen nicht vorhandene Hilfsmittel und nicht vorgeschriebene Kunstgriffe spontan ausgedacht werden.

Szymanski (*117*) fertigte auf einem Brett einen Schneckengang an, der mit kleinen Kieselsteinchen gefüllt wurde; dann gab er den Prüflingen (5—10jährigen Kindern) einen kleinen Besen in die Hand mit der Aufforderung, den Gang auszukehren. Die jüngeren Kinder handelten meist irrationell, indem sie am Ausgang der Windung begannen und nach innen hin kehrten; mit steigendem Alter nahm die Zahl der einsichtig handelnden, d. h. vom Mittelpunkt nach außen hin Arbeitenden zu. Der Altersfortschritt ist besonders stark von 8- zu 9jährigen (40% und 76% richtiger Fälle).

Rossolimo (*99*) schlägt einige der bekannten technischen Geduld- und Scharfsinnsaufgaben vor, z. B.: einen Metallring, der in eine mehrfach gewundene Spirale verschränkt ist, durch Drehungen und Gegeneinanderverschiebungen aus dieser herauszubringen (MS. 65—67). Einfacher, aber gerade zur Prüfung des eigentlichen Intelligenzanteils sehr passend, ist der Test von Lipmann und Stolzenberg (*201, 238*) (MS. 69): vorgelegt werden ein Schloß und zwölf Schlüssel, von denen zwei passen, sechs den passenden ähnlich und vier ganz abweichend sind; die zwei passenden Schlüssel sollen gefunden werden. Der Unintelligente wird mechanisch Schlüssel auf Schlüssel probieren; der Intelligente wird von vornherein die offenbar unmöglichen Schlüssel ausscheiden und dann die Erprobung mit dem wahrscheinlichsten beginnen.

Dies Beispiel zeigt bereits, daß wir es hier nicht mehr nur mit Fragen der technischen Eignung zu tun haben, sondern mit dem, was wir früher ganz allgemein „praktische Intelligenz" nannten (s. S. 23). Daß deren Prüfung in ganz anderem Maße als bisher in Angriff genommen werden muß, ist schon an früheren Stellen betont worden. Der Weg dazu ist uns auch bereits vorgezeichnet, und zwar durch die Affenversuche von Köhler (*26*) (s. S. 34). K. prüfte das Verhalten der Affen, wenn sie ein ihnen natürliches Verlangen (insbesondere nach Nahrung) nicht auf die angeborene oder eingeübte Weise befriedigen konnten, also zur Auffindung neuer Mittel und Kunstgriffe gezwungen waren. Im Anschluß an dies Verfahren hat Lipmann (*29*) für kleine Kinder analoge Aufgaben vorgeschlagen (MS. 68—69 und 247), die aber bisher noch der Erprobung harren. Es handelt sich stets darum, daß ein begehrter Gegenstand (Spielzeug, Näscherei), sichtbar, aber nicht ohne weiteres erreichbar ist; festgestellt wird, ob der Prüfling „sich zu helfen weiß", geeignete Hilfsmittel, z. B. einen Stock,

eine Schnur, eine Fußbank benutzt oder gar erst herbeiholt, sie zweckmäßig anwendet, mehrere wenn nötig verbindet usw. Die Schwierigkeit der Aufgabe kann hier beliebig abgestuft werden, und es ist zu erwarten, daß auf diesem Wege schließlich auch Methoden zur Prüfung der praktischen I. der höheren Jugendjahre gefunden werden[1]).

6. Aufmerksamkeits- und Konzentrationstests (MS. 70—73).

Was oben (S. 74) über die Beziehung der Aufmerksamkeit zur Merkfähigkeit gesagt wurde, gilt allgemein für ihre Beziehung zu intelligenten Handlungen überhaupt. Die Aufmerksamkeit an sich ist nicht ein wesentlicher Bestandteil der I. selbst, wohl aber eine unumgängliche Vorbedingung zu ihrer Betätigung. Eingeübte und mechanisch gewordene Verrichtungen können unter Umständen mit geringer Beteiligung der Aufmerksamkeit vollzogen werden; aber die Neueinstellung auf neue Anforderungen, die das Wesen der I. ausmacht, verlangt die Entfaltung geistiger Energie und ihre Zuspitzung auf das eine Ziel, d. h. Konzentration der Aufmerksamkeit. Die analytischen und synthetischen Aufgaben der I. setzen ferner voraus, daß man Mehreres zugleich im Bewußtsein habe, um es zu unterscheiden oder zu verbinden: also Distribution der Aufmerksamkeit. Endlich fordern die meisten I.-Aufgaben, daß die determinierende Tendenz des Zieles längere Zeit hindurch aufrechterhalten bleibe, um alle Teilphasen und einzelne Hilfsmittel zu bestimmen: also Ausdauer der Aufmerksamkeit. Eine geistige Begabung mag der Anlage nach noch so hochwertig sein — steht ihr nicht eine Aufmerksamkeitsbegabung von der nötigen Energie, Weite und Dauer zur Seite, dann wird sie nie aus bloßer potentieller Fähigkeit zu realer Fertigkeit werden. Deshalb ist es gerechtfertigt, daß man in die Serien für IP. auch Aufmerksamkeitstests eingestreut hat — nur darf man nicht (wozu Binet wohl zeitweilig neigte) glauben, damit die I. selbst zu treffen. Und es ist zu fordern, daß diese Aufmerksamkeitstests möglichst rein die Aufmerksamkeit prüfen, ohne zugleich noch besondere intellektuelle Schwierigkeiten zu bereiten.

[1]) Für die Untersuchung erwachsener Schwachsinniger ist die Methode soeben von Erich Stern *(111)* erfolgreich angewandt worden. Weitere Ausarbeitungen, die auch zur Prüfung normaler Erwachsener geeignet sind, finden zur Zeit im Hallenser Institut für praktische Psychologie statt. (Nach privater Mitteilung von Dr. Giese.)

In diesem Sinne ist der bekannte Bourdon-Versuch (*52*) empfehlenswert (MS. 70—71). Die Aufgabe, in einem vorgelegten Test bestimmte Buchstaben, z. B. alle a, e und r zu durchstreichen, stellt weder an Verständnis noch an Überlegung irgend welche Ansprüche, sondern fordert lediglich dauerndes, nie erlahmendes Aufmerken.

Weniger geeignet scheint der Test von Moede-Piorkowski (*206*), der die gleichzeitige Bearbeitung zweier disparater Aufgaben verlangt (Auffassen einer vorgelesenen Geschichte, während zugleich multipliziert wird, MS. 71—72). Die hier geprüfte Teilbarkeit der Aufmerksamkeit mag für manche Sonderberufe (z. B. für den Lenkerberuf) von Wichtigkeit sein; für die Schultätigkeit und für das allgemeine praktische Leben kommt sie kaum je in Betracht. Gewiß verlangen geistige Aufgaben oft genug das gleichzeitige Beachten von Verschiedenem; dann stehen aber diese verschiedenen Elemente unter dem gemeinsamen Gesichtspunkt des Endzieles und sind nicht völlig disparat zueinander, wie in jenem Test

Als „Aufmerksamkeitstest" empfiehlt neuerdings der Leipziger Lehrerverein (*197*) das Rechnen in einem nichtdekadischen System (MS. 235—236). Der Test ist sehr wertvoll, um bestimmte Seiten der mathematischen Begabung zu prüfen. Aber eben diese Abhängigkeit des Ergebnisses von einer Sonderfähigkeit scheint mir zu bewirken, daß der Test über die Aufmerksamkeitsleistung der Prüflinge keine eindeutigen Ergebnisse liefern kann.

7. Suggestibilität (MS. 73—75).

Auch die Suggestionstests gehören zu denjenigen, die, ursprünglich für andere psychologische Zwecke bestimmt, erst nachträglich für Zwecke der IP. nutzbar gemacht wurden.

Binet (*2*) bot auf Blättern, die er nacheinander vorlegte, verschiedene Linienpaare dar und ließ die Länge je zweier nebeneinander stehender Linien vergleichen (MS. 74). In den ersten drei Blättern war die rechtsstehende Linie etwas länger als die linksstehende; es wurde geprüft, ob dadurch eine Einstellung entstand, die suggestiv weiter wirkte, indem bei den folgenden Blättern mit objektiv gleich langen Linien nun auch noch die rechte Linie als die längere bezeichnet wurde.

Moede und Piorkowski (*206*) verwandten nach dem Vorbild der bekannten Aussageversuche Bilder, über deren Inhalt sie nach Wegnahme Suggestivfragen stellten (MS. 75).

Haben wir nun ein Recht, in dem Grade der Suggestibilität zugleich ein (negatives) Kennzeichen für den Grad der I. zu sehen? — Insofern gewiß, als der suggestible Mensch eine gewisse Unselbständigkeit des geistigen Verhaltens zeigt. Dort, wo von ihm eine eigene, spontane Stellungnahme verlangt wird, begnügt er sich damit, eine solche von anderwärts zu übernehmen — sei es,

daß er seine eigene frühere Stellungnahme automatisch fortsetzt (Binets Liniensuggestion), sei es, daß er die bei einem anderen vermutete Stellungnahme zu seiner eigenen macht (Eingehen auf Suggestivfragen beim Bildversuch). Mit dieser Scheu vor selbständiger Entscheidung verbindet sich dann eine mangelnde Kritik; denn der Mensch, der einer Suggestion unterliegt, glaubt ja aus Eigenem zu schöpfen, eine wirkliche Wahrnehmung, eine echte Erinnerung zu bringen; er verwechselt den suggerierten Inhalt mit Selbsterlebtem.

Aber damit ist die psychologische Charakteristik nicht erschöpft; denn die Suggestibilität ist überhaupt keine bloß intellektuelle Eigenschaft, sondern stark durch Faktoren des Gemüts- und Willenslebens mitbestimmt. Die Unselbständigkeit ist nicht nur eine solche des Denkens, sondern zugleich ein Bedürfnis, sich Autoritäten unterzuordnen, eine Scheu vor Konflikten; darum sagen schüchterne Menschen auf Suggestivfragen leicht „ja" und lassen dann erst, gleichsam sekundär, der gefühlsmäßigen die gedankliche Bejahung folgen. Aber auch einen ganz anderen Gemütsursprung kann die Suggestibilität haben, nämlich aus dem Ehrgeiz. Das Bedürfnis, anderen in gewissen Leistungen gleichzutun, führt dazu, daß man etwas zu sehen, zu hören, zu verstehen, zu wissen behauptet, was ein anderer einem vor-sieht, -hört, -versteht, -weiß — und man glaubt schließlich selbst, es wahrzunehmen oder zu wissen.

Dies Zusammenwirken so verschiedenartiger psychologischer Faktoren hat zur Folge, daß zwischen Suggestibilität und I. keine eindeutige Korrelation besteht; und in der Tat zeigen auch die experimentellen Befunde von Kosog u. a. (*76*), daß die begabteren Schüler einer Klasse nicht etwa einen stärkeren Widerstand gegen Suggestionen aufbringen als unbegabtere (MS. 75). Nur so viel darf man sagen, daß bei einem Menschen von hoher Suggestibilität die objektiv-kritische Seite der I. nicht stark entwickelt sein kann. Andere mehr subjektive und synthetische I.-Funktionen mögen unabhängig davon gut ausgebildet sein.

Eine sehr deutliche Beziehung hat dagegen die Suggestibilität zum Alter; das ist durch die Untersuchungen von mir selbst (*40*) über die Wirkung von Suggestivfragen festgestellt, später von Lipmann (*81*) bestätigt worden (MS. 74). Wenn nach meinen Befunden 7jährige in 50% aller Fälle, 14jährige dagegen nur noch in 20% den suggestiven Fragen über den Inhalt eines eben gesehenen Bildes erlagen, so spricht sich darin ohne Zweifel die Tat-

sache aus, daß die Spontaneität und Selbständigkeit des psychischen Verhaltens — sowohl in geistiger wie gemütlicher Hinsicht — mit den Jahren stark wächst.

8. Prüfung der Sprachbeherrschung (MS. 75—79).

Mit dieser Gruppe betreten wir das große Gebiet der sprachlichen Tests. Sie alle verlangen sprachliche Leistungen gewisser [illegible]len aus diesen auf die beteiligten Denkleistungen [illegible]nd zwar liegt entweder der Nachdruck in der sprachlichen Leistung als solcher — Prüfung der „Sprachbeherrschung" — oder aber die Aufgabe fordert eine Denktätigkeit, deren Material in sprachlichem Inhalt besteht und zu deren Vollzug und Äußerung sprachliche Akte als Mittel nötig sind.

Daß wir die Herrschaft eines Menschen über die Sprache feststellen, ist gewiß für eine Reihe theoretischer und praktischer Aufgaben von Wichtigkeit; so sind ja manche Berufe ohne eine bedeutende Sprachgewandtheit gar nicht denkbar. Aber Sprachgewandtheit darf nicht ohne weiteres mit I. gleichbedeutend gebraucht werden. Es gibt zweifellos eine sprachliche Sonderbegabung, bei der ein vorzügliches Wort- und Phrasengedächtnis, feines Sprachgefühl, sowie auch eine ästhetische Freude am sprachlichen Eindruck und Ausdruck eine weit größere Rolle spielen als die allgemeine I. So hat u. a. die Fähigkeit zu reimen, die Binet (*2, 3*) in seinem Staffelsystem prüft (MS. 76), ganz überwiegend mit dieser Sonderbegabung zu tun. Umgekehrt braucht sprachliche Schwerfälligkeit durchaus nicht immer durch geistige Stumpfheit bedingt zu sein; sie kann vielmehr auf eigentümlichen Hemmungen beruhen, die die Umsetzung der geistigen Inhalte in sprachlichen Ausdruck erschweren. Solche Hemmungen sind bei manchen mehr intellektueller Natur, indem die geistigen Inhalte selbst sich gegen sprachliche Formulierung sträuben (z. B. beim Mathematiker oder beim intuitiven Künstler), und es erfordert eine gewisse Vergewaltigung, sie in Sprachliches umzuformen. Bei anderen sind sie mehr emotioneller Natur: das seelische Innenleben scheut sich davor, sich nach außen darzustellen und dadurch zu entäußern.

Durch solche Betrachtung fällt wieder von einer neuen Seite her Licht auf die verwickelten Beziehungen zwischen Sprechen und Denken, die unter anderen Gesichtspunkten so vielfältig erörtert worden sind. Jedenfalls ergibt sich aus der Vieldeutigkeit dieser

Zusammenhänge, daß die bloße Feststellung der Sprachbeherrschung als solcher von allen sprachlichen Tests den unsichersten Symptomwert für die I. des Prüflings hat.

Hierher gehören die „Wortschatzprüfungen", die in zwei verschiedenen Formen auftreten: als Prüfung der Wortfindung und des Wortverständnisses.

Binet (*2, 3*) prüft den frei verfügbaren Wortschatz, indem er auffordert, beliebige Worte, die einem einfallen, zu sagen, zwar so zahlreich und so schnell wie möglich (MS. [redacted] Minuten sind dafür Zeit gelassen; 12jährige Kinder [redacted] auf mindestens 60 Wörter kommen. Die meisten Kinder fassen die Aufforderung so auf, daß sie Hauptwörter nennen sollen, und halten sich vor allem an die im Zimmer sichtbaren Gegenstände. Bei dieser Selbstbeschränkung auf das Wahrnehmbare ist freilich der Wortvorrat bald erschöpft, und nur schwer wird durch immer weiteres Drängen des Prüfers die Zahl der gefundenen Wörter vermehrt. Leichter haben es schon diejenigen Kinder, die sich von dem sinnlich Gegenwärtigen freimachen und Gegenstände der Straße, des Hauses, Verwandtschaftsnamen usw. nennen. Der entscheidende Fortschritt aber — und dieser hat wohl in der Tat etwas mit der I. zu tun — liegt darin, daß die Kinder ihre Aufzählung unter Gesichtspunkte bringen. Sie sind vielleicht zufällig auf „Pferd" gekommen und schließen nun die Reihe der anderen ihnen bekannten Haustiere an; die Erwähnung der Mutter eröffnet die Reihe der Verwandtschaftsnamen usw. Hier werden offenbar synthetische Ordnungsgrundsätze sichtbar, die vom Kinde spontan gefunden sind. Trotz dieser Beziehung zur I. ist der Test aber nicht zu empfehlen. Die Aufforderung, beliebige Worte zu nennen, stellt das Kind unter durchaus unnatürliche Bedingungen; denn sein Sprechen vollzieht sich nicht in einzelnen Worten, sondern in zusammenhängenden, seinem inneren Erleben entsprechenden Sätzen: und so sind zuweilen recht sprachgewandte Kinder bei diesem Test wortkarg. Außerdem hat der Test infolge des durch mehrere Minuten sich hinziehenden Drängens und Lauerns für Prüfer und Prüfling etwas Peinigendes.

Das Wortverständnis prüft Terman (*118*), indem er feststellt, welche Worte einer Darbietungsreihe dem Kinde bekannt sind (MS. 77—78). Die Reihe zeigt aufsteigende Schwierigkeit[1]).

[1]) Das Verfahren erinnert an die „Analysen des kindlichen Gedankenkreises", wie sie Hartmann u. a. angestellt haben.

Aber auch hier kommen Bedenken über Bedenken. Erstens spielt bei der Kenntnis vieler Worte das eingelernte Wissen und die Anregung des Hauses eine größere Rolle als die Intelligenz. Sodann ist die Bekanntschaft des Kindes mit dem Wort nur dadurch festzustellen, daß man eine Erklärung der Wortbedeutung geben läßt; daher vermischt sich die Methode mit der Prüfung der Definitionsfähigkeit, die doch eine ganz andere seelische Funktion ist; denn viele Kinder kennen manche Ausdrücke recht gut und wissen sie auch richtig anzuwenden, ohne imstande zu sein, anderen von ihrer Bedeutung Rechenschaft zu geben [1]).

Weniger die Sprachbeherrschung als eine gewisse Kombinationsfähigkeit prüft der Test: aus den Buchstaben eines Wortes so viel andere Worte wie möglich zu bilden (Whipple [*123*]) (MS. 76). Über den Wert dieses Tests liegen meines Wissens in Deutschland noch keine Erfahrungen vor.

9. Assoziationen (MS. 79—83).

Wären die sogenannten Assoziationsversuche wirklich nichts anderes als Prüfungen des „assoziativen" Vorstellungsverlaufs, so hätten sie unter den I.-Prüfungsmethoden nichts zu suchen. Assoziationen im eigentlichen Sinne sind ja jene Vorstellungszusammenhänge, wie sie unabhängig von Absicht und Zielsetzung, rein durch die Mechanik früherer Verkettungen von selbst sich einstellen; die Betätigungen der I. aber sind dadurch gekennzeichnet, daß der Ablauf der geistigen Inhalte erstens neue, nicht mechanisch eingefahrene Wege geht, und daß er zweitens unter Zielen steht und durch sie bestimmt ist. Aber jenes reine Assoziieren läßt sich auch gar nicht experimentell fassen; denn schon die bloße Aufgabe: „Ich soll assoziieren" macht jenen Zustand der ziellosen Passivität unmöglich, in dem allein der bloße Assoziationsmechanismus spielt.

Die Assoziationsexperimente sind also in Wirklichkeit Denkaufgaben — freilich in verschiedenem Maße. Am geringsten ist der Denkanteil bei den sogenannten freien Assoziationsprüfungen: auf jedes zugerufene Wort soll mit dem ersten sich einstellenden Wort geantwortet werden; über die Beschaffenheit dieses Wortes und sein logisches Verhältnis zum Reizwort wird keinerlei Vorschrift gemacht. So ergiebig diese Methode zur Feststellung der

[1]) Vgl. hierzu Abschnitt 14 dieses Kapitels: Definieren (S. 104ff.).

verschiedenen Assoziationsformen, ferner für Ermüdungsuntersuchungen, für psychopathologische Untersuchungen, für Komplexforschung und Tatbestandsdiagnostik geworden ist, so gering ist die Ausbeute für das I.-Problem. Weder der Inhalt der Assoziationen noch das Tempo des Assoziierens steht, wie Winteler (*124*) und andere fanden, in eindeutiger Korrelation zur geistigen Allgemeinbegabung. Höchstens kann vielleicht der Seltenheitswert der Assoziation eine Rolle spielen. Es läßt sich massenstatistisch feststellen, daß zu bestimmten Reizworten gewisse Assoziationen am häufigsten auftreten (z. B. auf schwarz: weiß), andere seltener (schwarz: Neger). Bei Kindern mittlerer I. scheinen nun banale Assoziationen am häufigsten zu sein; starke Abweichungen von den üblichen Assoziationen findet man einerseits bei zurückgebliebenen Kindern (Roemer [*98*]), bei denen die selbstverständlichen Verbindungen fehlen und ganz abseitige und sinnlose vorwiegen, andererseits bei stark begabten Kindern, die eine gewisse Originalität des Vorstellungsverlaufs besitzen.

Ganz anders steht es um die „gebundenen" Assoziationen, bei denen die zu findenden Worte ein vorgeschriebenes gedankliches Verhältnis zu den Reizworten besitzen müssen. Hiermit ist die geistige Tätigkeit unter einen Gesichtspunkt gestellt, der nun eine dauernde Denkeinstellung während der ganzen Versuchsreihe schafft und eine determinierende Tendenz auf jede einzelne Assoziation ausübt. Zwei Leistungen eigentlich intellektueller Natur werden also gefordert: einmal das Festhalten einer abstrakten Aufgabe bei fortwährend wechselnden konkreten Reizen; sodann die richtige Spezialisierung der Aufgabe auf den Einzelfall.

Die Methode ist von Sommer (*110*) und anderen Psychiatern an Geisteskranken, von Winteler (*124*) und Ries (*95*) an Schulkindern angewandt worden.

Winteler (MS. 81) verlangte bestimmte begriffliche Beziehungen: Unter-, Neben-, Überordnung und Entgegensetzung. Am deutlichsten treten die I.-Unterschiede der Prüflinge zutage bei der Aufgabe, zu den zugerufenen Wörtern (z. B.: Kuh, Veilchen, Montag, Schuh, Töpfer, Spiegel) den übergeordneten Begriff zu nennen. Nicht nur durch die Schnelligkeit und Richtigkeit der Reaktion waren die Intelligenteren überlegen, sondern auch durch die Schärfe der Begriffsbeziehung. Während nämlich weniger Intelligente dazu neigen, einen ganz allgemeinen Oberbegriff zu nennen (Kuh: *Tier;* Töpfer: *Mensch*), suchen die Intelligenteren

den gegebenen Begriff genauer zu charakterisieren durch Nennung der nächsthöheren Gattung (Kuh: *Haustier;* Töpfer: *Handwerker*) und entsprechen dabei unbewußt der Forderung, die die Logik an die begriffliche Definition stellt.

Ries (später auch Weigl [*122*]) (MS. 79—80) prüfte das funktionale Denken, indem als Assoziationsgesichtspunkt das Verhältnis von Ursache und Wirkung vorgeschrieben wurde. Zu Wörtern wie: Irrtum, Friede, Bestechung, Wärme, Ernährung, Nässe, Schiffbruch usw. waren andere zu finden, die die Wirkungen der Reizwörter ausdrückten. Dieser Test hatte besonders günstige Ergebnisse; seine Korrelation mit der I.-Schätzung durch die Lehrer betrug in verschiedenen Klassen 0,85—0,95, war also außerordentlich hoch. Es ist dies verständlich, weil das natürliche Denken sich viel mehr in funktionalen Bahnen als in solchen der Begriffsordnung bewegt. Der Test hat eine nahe Verwandtschaft mit dem ebenfalls von Ries angegebenen Test, der die Merkfähigkeit für ursächlich verknüpfte Wortpaare prüfte (S. 76); dort wurde das Verstehen gegebener, hier wird das Finden aufgegebener Kausalzusammenhänge verlangt.

Nicht mit Begriffspaaren, sondern mit einer ganzen Kette hat es der Assoziationstest von Moede-Piorkowski (*206*) zu tun (MS. 80). Zu einem einzelnen Wort (z. B. Baum, Maschine) sollen möglichst viele damit in Zusammenhang stehende Worte assoziiert werden. Hier soll also der Reichtum eines einzelnen Vorstellungskomplexes an innerlich zusammengehörigen Elementen und die Verfügbarkeit dieser Elemente auf einen gegebenen Anstoß hin geprüft werden. Der Test scheint für ältere und begabte Kinder geeignet zu sein, doch ist seine exakte Bewertung schwierig.

10. Bildbetrachtung (MS. 83—97).

Die Darbietung von Bildern ist eine der beliebtesten und ergiebigsten Methoden der angewandten Psychologie geworden. Dies liegt an der Zwischenstellung des Bildes zwischen dem strengen Laboratoriumsversuch, dessen Exaktheit durch Lebensfremdheit erkauft wird, und der unmittelbaren Lebenswirklichkeit, die sich in ihrer Verwickeltheit und ständigen Veränderung jeder exakten Verwertung entzieht. Das Bild hat Lebensnähe, stellt Ausschnitte der Wirklichkeit dar, nicht nur künstlich isolierte Wahrnehmungselemente; und es ist zugleich experimentell verwertbar, kann dem

Problem entsprechend ausgesucht oder hergestellt und allen Prüflingen unter genau vergleichbaren Bedingungen vorgelegt werden. Die Ergebnisse lassen sich an dem ja stets bequem zur Verfügung stehenden Objekt kontrollieren und unter den verschiedensten Gesichtspunkten psychologisch verarbeiten. So wurde das Bildexperiment herangezogen zur Untersuchung der Beobachtungs- und Erinnerungsfähigkeit, der Suggestibilität und der Phantasie, der Raumauffassung, des ästhetischen Verständnisses, der Einfühlung und schließlich auch der I.

Für den Zweck der I.-Prüfung sind Bilder deswegen so wertvoll, weil sie der spontanen geistigen Tätigkeit Spielraum geben. Die Aufgabe, den Inhalt eines Bildes oder einer Bilderserie sprachlich darzustellen, gibt niemals eine festgebundene Marschroute, sondern verlangt selbständige analytische und synthetische Akte. Jedes Bild, auch das einfachste, bietet unbestimmt viele Möglichkeiten des Erwähnbaren, an Einzelelementen, an Beziehungen dieser Elemente zueinander, an Verknüpfungen, die über den Bildinhalt hinausweisen auf das, was vorher und nachher liegt, an subjektiven Stellungnahmen; und nun ist es dem Prüfling selbst anheimgestellt, wie er einerseits — analytisch — die Auslese trifft, und wie er andererseits — synthetisch — Zusammenhänge erkennt und herstellt. Je weniger eng hier durch die Versuchsanordnung die Eigentätigkeit des Prüflings begrenzt ist, um so mehr kann er seine spontane I.-Richtung betätigen; um so mehr nähert sich der Bildversuch der natürlichen Lebenswirklichkeit an. Darum sind für unsere Zwecke nicht so sehr die Verhörsmethoden wichtig, die jeden zu beschreibenden Punkt genau umgrenzen, sondern die Berichtsmethoden, die eine zusammenhängende — mündliche oder schriftliche — Darstellung des Bildinhalts fordern. Wie sich in der Auslese des Bildstoffes die verschiedenen Entwicklungsstadien der I. darstellen, ist an anderer Stelle (S. 39) ausgeführt worden.

Ein weiterer methodischer Unterschied ergibt sich dadurch, das man das Bild entweder während oder nach der Darbietung beschreiben läßt. Im ersten Fall, dem „Beobachtungsversuch", wird rein die Auffassung und das Verständnis geprüft, im zweiten Fall, dem „Aussageversuch", kommt noch die Merkfähigkeit und die Erinnerungstreue dazu. Bei Bildern, deren Stoff schon an sich gewisse Schwierigkeiten bietet, ist das erste Verfahren angemessen; bei leichteren Bildern oder in Fällen, in denen die Aufgabe erschwert werden soll (Auslese der Begabten, Prüfung älterer Jahr-

gänge), wendet man besser die Aussagemethode an. Bei ihr besteht nämlich der Zwang, schon während der Betrachtung des Bildes die Aufmerksamkeit auf die wesentlichen Züge zu konzentrieren, da man nicht alles behalten kann. Die Fähigkeit, Wesentliches und Unwesentliches voneinander zu scheiden, ist aber ein wichtiges Unterscheidungsmerkmal der I.

Der Darbietung von Einzelbildern hat sich in neuerer Zeit die Verwertung von Bilderfolgen zugesellt, die nach unseren Breslauer und Hamburger Erfahrungen recht brauchbar sind (MS. 94—95). In solchen Bilderfolgen (wie sie z. B. in vielen Münchner Bilderbogen vorliegen) wird eine zusammenhängende Handlung durch eine Reihe von Phasen dargestellt; natürlich muß ein etwa vorhandener verbindender Text vor dem Versuch entfernt werden. Die Aufgabe für den Prüfling ist hier vor allem kombinatorischer Natur: er muß aus den Einzelphasen den Zusammenhang konstruieren, also die nicht dargestellten Übergänge richtig ergänzen. Berichte über solche Bilderfolgen werden jetzt meist in schriftlicher Form gefordert; sie haben sich nicht nur für psychologische Prüfungszwecke bewährt, sondern werden auch von Pädagogen als eine neue reizvolle Form des Aufsatzes begrüßt. Die Ausbeute solcher Aufsätze an psychologischen Befunden ist außerordentlich vielseitig: Beobachtungsfähigkeit und Kombinationsgabe, bei dieser wieder der vorwiegende Anteil von logischer Ergänzung oder von intuitiver Phantasie, die Erfassung des Wesentlichen, die Fähigkeit der Einfühlung in die verschiedenen Erlebnisse und die typischen Eigenschaften der dargestellten Personen, Empfänglichkeit für die ästhetischen Werte oder für den Humor der Bilder und manches andere läßt sich daraus entnehmen und individualisierend oder vergleichend verwerten.

Eine aussichtsvolle Erweiterung der Bilderbogenmethode ist neuerdings durch Penkert (*216*) und Susanne Engelmann (*187a*) ausgeprobt worden: man verlangt über denselben Bilderbogen verschiedene Darstellungen nach jeweils anderer Instruktion, z. B.: 1. „so anschaulich und lebendig wie möglich", 2. „so knapp wie möglich", 3. „Vorgeschichte und Nachgeschichte", 4. „so, wie man den Bilderbogen einem fünfjährigen Kind erzählen würde". Das bessere oder schlechtere Gelingen bei der einen oder anderen Aufgabestellung gibt Auskunft über den I.-Typ des Prüflings.

Nicht geringe methodische Schwierigkeiten bereitet die Bearbeitung der Ergebnisse bei Bildversuchen. Je weniger gebunden die Aufgabe, je komplexer die beteiligten seelischen Vorgänge, um so schwerer ist es, eine exakte Zergliederung und Bewertung der

Leistungen durchzuführen. Es wird nötig sein, jede einzelne Leistung nach einer ganzen Reihe von Gesichtspunkten zu analysieren, zugleich aber auch den Gesamteindruck und die hierdurch nahegelegte Zugehörigkeit zu einem I.-Grad oder -Typ nicht aus den Augen zu verlieren. Wie weit eine solche analytische Bearbeitung gehen kann, dafür bietet die Hamburgische Prüfung mit dem Bilderbogen „Das Wiedersehen" ein Beispiel (MS. 96).

Um festzustellen, ob das Wesentliche erfaßt sei, kann man wiederum ein analytisches und ein synthetisches Verfahren einschlagen. Man entwirft ein Schema der „Hauptpunkte", die das logische Gerippe der Darstellung bilden müssen, wenn der wesentliche Zusammenhang überhaupt richtig zur Darstellung gelangen soll, und untersucht, ob diese Punkte in der Niederschrift enthalten sind. Andererseits prüft man, ob die „Pointe" des Gesamtinhalts von dem Prüfling spontan zu selbständigem Ausdruck gebracht sei, oder auch, ob er auf die Aufforderung, selbst eine Überschrift zu seinem Aufsatz zu finden, einen wirklich treffenden, d. h. das Wesentliche hervorhebenden Titel wähle.

11. Erkennungstests (MS. 98—106).

Auch hier handelt es sich durchweg um die Auffassung bildlicher Darstellungen, aber um solche einfacher Gegenstände, die unvollständig oder undeutlich wiedergegeben sind. Der Prüfling hat also eine Ergänzungsleistung sehr elementarer Art zu vollziehen: wahrgenommene Linien, welche Umrisse oder Bruchstücke darstellen, sollen zur Gegenstandswahrnehmung vervollständigt werden. Die Tests sind wegen ihres elementaren Charakters besonders für jüngere Kinder geeignet.

Der Ergänzungsvorgang beim Kinde geht nun nicht etwa immer so vor sich, daß das Fehlende als solches bemerkt und in der Vorstellung hinzugefügt würde. Das Ursprünglichere ist es, daß das Kind mit dem vorhandenen Bruchstück unmittelbar eine Gegenstandsvorstellung assoziiert, ohne sich darüber klar zu werden, worin diese Vorstellung von der tatsächlich gegebenen Wahrnehmung abweicht, bzw. über sie hinausgeht. Selbst wenn das Kind nicht sofort zur Erkennung gelangt, sondern nachdenken und suchen muß, braucht ihm das im Bild Fehlende nicht zum Bewußtsein zu kommen.

Das Bemerken des Fehlenden oder das Vermissen ist erst

eine höhere Form der Erkenntnis; ihr ist ein besonderer Test von Binet (*3, 4*) gewidmet (MS. 98—99), in dem unvollständige Menschenfiguren und Gesichter gezeigt werden und vom Prüfling die Angabe der fehlenden Körperteile verlangt wird.

Die exakteste Ausgestaltung hat der Erkennungstest in den sogenannten Heilbronner-Serien (*67, 120*) erfahren (MS. 101 bis 103). Ein Gegenstand wird in einer Reihe immer vollständiger werdender Zeichnungen hintereinander gezeigt, so daß stets nur ein Bild auf einmal sichtbar ist. Untersucht wird erstens, bei welchem Grade der Vollständigkeit Erkennung erfolgt, und zweitens, ob die kleinen Änderungen von Bild zu Bild bemerkt werden. Das Ergänzen der zuerst nur sehr dürftigen Bruchstücke oder Umrisse zu einem sinnvollen Gegenstand ist eine ausgesprochen synthetische I.-Leistung; die Unterscheidung der Bilder in bezug auf die jedesmal neu hinzukommenden Elemente eine analytische Tätigkeit. Dieser Test ist neuerdings von G. und A. Schober (*108*) in Hamburg durch Schaffung zweckmäßiger Bilder (die Heilbronnerschen haben manche Fehler und sind für deutsche Verhältnisse zum Teil nicht passend) und durch Eichung für verschiedene Altersstufen ausgestaltet worden, so daß er jetzt als recht geeigneter Test für die frühe Kindheit (etwa bis zum 8. Jahre) gelten kann (MS. 103 bis 106).

12. Verhalten zu vorgestellten Situationen (MS. 106—125).

Unter diesem bisher nicht üblichen Titel wird eine Reihe von Aufgaben zusammengefaßt, die alle das Gemeinsame haben, daß sie ein fiktives Verhalten erfordern. Es ist diejenige Stelle, wo sich theoretische und praktische I. am nächsten kommen, ohne sich aber doch zu verschmelzen. Stets handelt es sich um eine praktische Lebenslage; es soll eine Handlung ausgeführt oder eine Handlungsweise gewertet werden — aber die Situation ist keine wirkliche, sondern eine bloß gedachte; und der Prüfling soll nicht tatsächlich handeln, sondern ein potentielles Handeln vorstellen und beschreiben.

Die „Fiktionsfähigkeit" ist zweifellos eine sehr bedeutsame Seite der I. Der Prüfling muß sich von der rein realistischen Verstricktheit des wirklichen Lebens frei machen und etwas Nichtseiendes als seiend denken können; er muß ferner imstande sein, Situationen, die bei wirklichem Auftreten eine vielleicht instink-

tive und unbewußte Reaktion hervorrufen, sich innerlich klar vorzustellen, über sie nachzudenken und seine mögliche Reaktion, ja vielleicht die zu ihr führenden Motive ins Bewußtsein zu heben.

Solche Fähigkeiten gehören ja zum Teil in das Gebiet der Phantasie; aber da es sich nicht um beliebiges Vorstellen von Nichtseiendem handelt, sondern um Feststellung des zweckmäßigsten Verhaltens zu der fingierten Situation, um Beurteilung von Möglichkeiten und Wahrscheinlichkeiten, um Ziehung naheliegender Folgerungen, so haben wir es doch in erster Reihe mit einer I.-Leistung zu tun. Und zwar ist diese Fiktionshandlung von besonderer praktischer Wichtigkeit; denn sie löst das praktische Handeln von den Augenblicksbedingungen der Gegenwart los und befähigt den Menschen, künftige Möglichkeiten des eigenen Handelns denkend vorwegzunehmen oder sich in fremdes Handeln hineinzuversetzen. Hier liegt ein typischer Unterschied der menschlichen und der tierischen I. vor.

Das fiktive Verhalten gliedert sich nach zwei sich kreuzenden Gesichtspunkten, einmal nach der Art der Stellungnahme, sodann nach der Fremdbezogenheit oder Ichbezogenheit der Situation.

Die Stellungnahme kann entweder intellektuell oder gefühlsmäßig und wertend sein.

Im ersten Falle gilt es, fingierte Situationen zu verstehen und in ihren möglichen oder wahrscheinlichen Ursachen und Wirkungen zu beurteilen. Handelt es sich hierbei nicht um eigene Lebenslagen, sondern um solche anderer Menschen, so ist der Widerstreit zwischen Fiktion und Wirklichkeit am geringsten. Denn wenn wir eine fremde Lage nur kühl verstandesmäßig auffassen, so ist es schließlich von geringem Einfluß, ob diese Situation tatsächlich vorliegt oder nur vorgestellt ist. Das Wesentlichste ist hier die Fähigkeit, auf Grund bloßer Beschreibung die Situation überhaupt genügend klar vorzustellen, um die nötigen Operationen des Verstehens und Beurteilens daran zu knüpfen.

Einfache Aufgaben dieser Art sind die von Binet und seinen Nachfolgern angewandten kleinen Geschichten (MS. 106—107), die den Lokalnotizen der Zeitungen nachgebildet sind, deren Abschluß aber vom Prüfling selbst gefunden werden muß, z. B.: „Ein Mann, der im Wald spazieren ging, blieb plötzlich erschrocken stehen und lief zu dem nächsten Polizisten, um ihm mitzuteilen, daß er soeben an einem Baumzweig hängen sah . . . Was hatte er gesehen?“

Moede und Piorkowski (*206*) gestalten den Test für ältere, besonders begabte Kinder in der Weise, daß sie ein interessantes Ereignis (einen Unfall, ein Verbrechen) schildern und die verschiedenen Möglichkeiten der Ursachen oder Folgen bald nach dem Grade der Wahrscheinlichkeit, bald nach dem der Zweckmäßigkeit beurteilen lassen (MS. 108—110). Hierfür ist natürlich ein recht hoher Grad von souveräner geistiger Beherrschung des fingierten Stoffes notwendig.

Das psychologische Bild wird ein anderes, wenn es sich um eigene Lebenslagen handelt. Hierher gehören die ebenfalls von Binet und seinen Nachfolgern ausgebildeten „Verstandesfragen" (MS. 111—115), welche meist die Form haben: „Was soll man tun (oder: Was würdest du tun), wenn . . .?" Selbst wenn die Frage in der unpersönlichen „man"-Form gestellt ist, hat sie doch stets den Sinn — und wird auch immer so aufgefaßt —, daß der Gefragte sich selbst im Geiste in die angegebene Lage versetzen und daraus für sich die Folgerungen ziehen soll. Darum sind solche Fragen nur dann brauchbar, wenn sie kindgemäße Situationen vorlegen, z. B.: „Was muß man tun, wenn man etwas zerbrochen hat, was einem nicht gehört?" „. . . wenn man etwas bestellen soll und es unterwegs vergißt?" usw. Recht unkindlich und darum ungeeignet ist die von Binet für 11—12jährige Kinder angewandte Frage: „Denk mal, es fragt dich jemand nach deiner Meinung über einen anderen Menschen, den du nur wenig kennst; was würdest du dann sagen?"

Man hüte sich übrigens vor der Meinung, daß man mit diesen Fragen das tatsächliche praktische Verhalten des Kindes prüfe. Nehmen wir die oben erwähnte Frage: „Was muß man tun, wenn man etwas zerbrochen hat, was einem nicht gehört?" Ein Kind mag hierauf ganz richtig antworten: „man muß es ersetzen" oder „man muß um Entschuldigung bitten"; aber im Ernstfall wird vielleicht dasselbe Kind mit Selbstverständlichkeit fortlaufen und sich verstecken oder gar durch eine Ausrede die Sache von sich abzuwälzen suchen. Dabei braucht die andersartige Beantwortung der Frage nicht etwa auf bewußter Heuchelei zu beruhen; unsere wirklichen Triebfedern decken sich eben durchaus nicht mit den Reflexionsmotiven, die wir bei einer nur gedachten Situation zu erleben glauben.

Diese Erwägungen gelten in noch höherem Maße, wenn es

sich nicht um eine intellektuelle, sondern um eine gefühlsmäßige oder wertende Stellungnahme zur fingierten Situation handelt (MS. 115—125). Die Antworten auf Fragen wie: warum man das Stehlen für verboten halte[1])? oder: was man mit einem gestohlenen oder gefundenen Geldstück anfangen würde[2])? usw. geben uns niemals Auskunft über das tatsächliche Motivationsgetriebe und damit über die sittliche Beschaffenheit des Prüflings, sondern nur über seine Einsicht in die hierbei zu berücksichtigenden moralischen Gesichtspunkte. Die geistige Beherrschung dieser Gesichtspunkte ist lediglich eine Vorbedingung für das höhere moralische Verhalten; und insofern ist es von Bedeutung, ob das Verbot des Stehlens nur durch Angst vor Strafe, oder durch Rücksicht auf die Angehörigen, oder durch Anerkennung des Eigentumsbegriffs, oder durch Hinweis auf die Rechtssicherheit der Gemeinschaft begründet wird. Aber die Willensbeschaffenheit selber wird durch diese Aussagen nicht bekundet, selbst unter der — nicht immer zutreffenden — Voraussetzung, daß die Angaben mit subjektiver Aufrichtigkeit gemacht worden sind.

Deshalb prüfen solche Tests die Einsicht in moralische Wertmaßstäbe und die Fähigkeit, auf Grund dieser Maßstäbe Werturteile zu fällen, also intellektuelle Funktionen; und deshalb gehören sie in unser Gebiet der I.-Prüfungen, nicht aber in das einer „Moralitätsprüfung". Tests für eine solche gibt es zurzeit nicht und wird es vermutlich niemals geben; die Meinung, daß wir mit Hilfe von Tests je zur Aufstellung von „Stufen des Moralitätsalters", entsprechend den Stufen des „Intelligenzalters", kommen könnten, halte ich für utopisch. Aber auch schon jenes viel bescheidenere Ergebnis, daß wir etwas von der moralischen Einsicht und Urteilsfähigkeit des Prüflings erfahren, ist schon wertvoll genug; hängt doch z. B. die forensische und erziehliche Behandlung von jugendlichen Delinquenten oder von Fürsorgezöglingen in hohem Maße von dem Grade dieser Einsicht ab[3]).

Höhere Stufen dieses Tests treten dort auf, wo die Situation

---

[1]) Schäfer (*106*); Levy-Suhl (*79*).

[2]) Riebesell (*93*).

[3]) Die Bezeichnung „Gesinnungsprüfung" für diese und verwandte Tests scheint mir deshalb gleichfalls nicht recht zutreffend zu sein. Wir finden diese Benennung bei Dr. Hjalmar Sander (*105*), dessen Zusammenstellung und kritische Besprechung aller auf moralische Einsicht und Urteilsfähigkeit bezüglichen Prüfmittel im übrigen von hohem Werte ist.

eine ungewöhnliche Entwicklung zeigt oder wo mehrere Handlungen einem vergleichenden Werturteil unterzogen werden sollen. Verwickelte moralische Fälle legt Roth (*102*) vor (MS. 116—120), indem er schildert, wie Menschen im Widerstreit von Pflichten gehandelt haben, wozu dann der Prüfling Stellung zu nehmen hat; R. versuchte auf diese Weise den verschiedenen Altersstufen seiner Gymnasiasten bestimmte vorwiegende Gebiete der moralischen Wertung zuzuweisen.

Auch die von Moede-Piorkowski (*206*) benutzte Geschichte „Das große Los" (MS. 121—122) gehört hierher, wenn sie es auch nicht mit eigentlich sittlichen Problemen zu tun hat. Ein Mann spielt seit seiner Jugend in der Lotterie und gewinnt als ganz alter Mann das große Los; es sollen die Gefühle und Gedanken des greisen Gewinners geschildert werden. (Auch hier handelt es sich ja um den Konflikt zwischen dem äußeren Glückszufall und dem bitteren Groll über die Ironie des Schicksals, das ihm den Reichtum in den Schoß wirft, da er nichts mehr damit anfangen kann.)

Allen solchen Fällen gegenüber scheint es zwei typisch verschiedene Verhaltungsweisen zu geben: die kühle sachliche Beurteilung des Falles als eines fremden, und der Versuch, sich in den Handelnden oder Erlebenden hineinzuversetzen, unmittelbar durch „Einfühlung" oder mehr intellektuell durch „Eindenkung". Dort wird man dazu neigen, allgemeine Grundsätze (die auch weiter nichts als eingelernte Morallehren zu sein brauchen) an den Fall anzulegen — eine bloße Subsumtionshandlung —; hier wird man mehr die seelische Verfassung des Handelnden, die man in ihrer konkreten Besonderheit innerlich nacherlebt, zur Grundlage des Urteils machen. Das zweite Verhalten ist sicherlich das höhere; und die bisherigen Testbefunde zeigen auch, daß ein solches wirkliches Sich-Einleben in die seelischen Erregungen und Kämpfe von anderen erst von der Pubertät an und bei den Höherbefähigten möglich ist. (Über diese Fähigkeit der Einfühlung und Eindenkung vgl. auch S. 43.)

Die Aufgabe, moralische Probleme vergleichend zu beurteilen, gibt Jacobsohn (nach Fernald); er legt sieben verschiedene Vergehen und Verbrechen vor (*70*) und fordert vom Prüfling, daß er sie der Schwere nach ordne (MS. 123—125). Auch hier wird zwar die Absicht des Verfassers, das sittliche Fühlen festzustellen, nicht erreicht, wohl aber wiederum die Urteilsfähigkeit auf sittlichem Gebiet geprüft. Freilich müßten die Fälle für das rein intellektuelle Verständnis gleich leicht und dem Kinde

gleich zugänglich sein, was für die Jacobsohnschen Proben noch nicht gilt.

Umfassende Nachprüfungen, die Jacobsohn soeben mit diesem Test an Fürsorgezöglingen wie an Lyzeistinnen und gebildeten Erwachsenen anstellte (*70*a) lehrten u. a. folgendes. Unterschiede der sittlichen Urteilsfähigkeit ergaben sich einerseits aus der aufgestellten Reihenfolge der sieben Delikte, andererseits und vor allem aus den Begründungen, die die Prüflinge für ihre Anordnung gaben. Es zeigte sich, daß die niederen Intelligenzen vornehmlich nach dem angerichteten Schaden, also dem Erfolg, die höheren nach den zugrunde liegenden Handlungsmotiven urteilten.

13. Kritikfähigkeit (MS. 125—138).

Das Kritisieren ist eine höhere Stufe des Verstehens; der Mensch begnügt sich nicht damit, das Gegebene in seinem positiven Inhalt aufzufassen; er mißt es an einer im Bewußtsein bereitliegenden Norm, die sich in dem dargebotenen Stoff nicht befriedigt findet.

Die erste Phase der kritischen Tätigkeit ist rein negativer Art: irgend ein Element des vorliegenden Inhalts wirkt störend; der mehr gefühlsmäßigen Abneigung schließt sich eine intellektuelle Ablehnung, ein absprechendes Werturteil an: das ist falsch, häßlich, schlecht usw. Die zweite Phase besteht in der Reflexion: es wird festgestellt, was an der bemängelten Stelle falsch ist, vielleicht auch das Motiv der Verfehlung aufgedeckt. In der dritten Phase endlich wird die Kritik zur positiven Leistung der Berichtigung: an die Stelle der fehlerhaften Form wird die zutreffende gesetzt. Unter Umständen kann hier die Arbeit des Kritikers geradezu schöpferisch wirken, indem er neue Wege weist, die aus den bisherigen Unzulänglichkeiten hinauszuführen vermögen.

Dem Gegenstande nach kann sich die Kritik auf alle Gebiete des Daseins erstrecken, auf denen es überhaupt Wertunterschiede gibt. Das ethische und das ästhetische Wertgebiet lassen wir an dieser Stelle beiseite; soweit sie in Kritiktests zur Geltung kommen, werden sie an anderer Stelle behandelt[1]). Hier gehen uns nur die intellektuellen Kritiken unter dem Gesichtspunkt richtig und falsch an; sie lassen sich wieder in logische und empirische gliedern, die freilich in concreto nicht immer scharf voneinander

[1]) So kommt bei den Tests des Abschnittes 12 die moralische Kritik unter dem Gesichtspunkt gut und schlecht zur Geltung, bei dem Vergleich von Gesichtern S. 120, Anm. die ästhetische Kritik unter dem Gesichtspunkt schön und häßlich.

zu scheiden sind. Die logische Kritik geht auf Verstöße gegen den Sinn, die empirische Kritik auf Verstöße gegen Erfahrungsgesetze und gelernte Regeln. Ein logischer Unsinn liegt in den Sätzen: „Ich habe drei Brüder, Paul, Ernst und ich", oder: „Die beiden Schwestern sehen sich sehr ähnlich: namentlich bei der einen war die Ähnlichkeit besonders auffallend." Empirische Unmöglichkeiten enthalten dagegen die Sätze: „Am ersten Tage, als die Wälder zu grünen begannen und der Fluß auftaute, nahm ich eine Sense und ging zum Ernten", oder: „Ein Arbeiter fiel vom Gerüst und brach ein Bein; um sofortige Pflege zu haben, lief er schnell ins Krankenhaus." Die empirischen Kritiken setzen demnach ein bestimmtes Wissen voraus, und eben deshalb sind sie für die Zwecke der experimentellen Prüfung weniger zu verwenden; denn oft genug mag das Versagen der Kritik nicht so sehr auf eigentlicher Kritikunfähigkeit beruhen als darauf, daß die zum Erkennen des Fehlers nötigen Kenntnisse nicht vorhanden oder wenigstens nicht in genügender Bereitschaft sind. Der Satz: „Im Tannenwald waren die Bäume wegen der frühen Jahreszeit noch ganz kahl" mag von einem Stadtkind, das kaum je in der freien Natur gelebt hat, unbesehen hingenommen werden, weil ihm die einschlägigen Erfahrungen fehlen.

Leider ist unser Schulbetrieb viel zu sehr auf diese empirische Kritik eingestellt, und zwar bei Lehrern und Schülern; man achtet vor allem darauf, ob gegen die gelernten Regeln verstoßen wird. Sehr bezeichnend ist in dieser Hinsicht der Ausfall einer Prüfung, die in Hamburg an 14- und 15jährigen Mädchen (Bewerberinnen um Aufnahme in ein Lehrerinnenseminar) angestellt wurde. Sie erhielten einen deutschen Text vorgelegt, in den zahlreiche Fehler, und zwar ebensowohl logische wie empirische, eingestreut waren. Während die letzteren, also Fehler gegen die Grammatik, den Satzbau und die Interpunktion, zu einem großen Prozentsatz bemerkt wurden, blieben die logischen Widersinnigkeiten zum großen Teil unbemerkt (MS. 131—134) (*216*).

Es wird in Zukunft bei der Gestaltung von Kritiktests möglichst darauf zu achten sein, daß zum Erkennen der Fehler nicht mehr Wissen vorausgesetzt wird, als bei den Kindern des betreffenden Alters mit Selbstverständlichkeit vorhanden ist. Das Wesentliche der zu bemängelnden Fehler muß auf logischem Gebiet liegen; nur dann können wir in der Kritikleistung wirklich ein Symptom der Intelligenz erblicken.

Des weiteren lassen sich die Kritikaufgaben scheiden in sprach-

lich-begriffliche, bildliche und technisch-kausale. Bisher sind ganz überwiegend Tests der ersten Form ausgebildet worden und auch unsere obigen Beispiele gehören durchweg dieser Gruppe an. Bildliche Widersinnigkeiten sind von Rossolimo (*99, 101*) in Vorschlag gebracht worden, z. B.: Eine Dame liest mit verbundenen Augen; in einer Landschaft fällt der Schatten der Bäume in der Richtung zur Sonne hin usw. (MS. 136—138). Die Ausbildung von technisch-kausalen Kritiktests ist erst von der Zukunft zu erwarten; mir liegen verschiedene Vorschläge vor, die aber noch nicht durchgeprüft sind. Es müßte z. B. in einem kleinen Modell oder auch im Bilde eine technische Konstruktion gezeigt werden, bei der die beabsichtigte Bewegung durch irgend eine Vorrichtung unmöglich gemacht wird; natürlich darf es sich nur um ganz einfache, keine technischen Kenntnisse voraussetzenden Mechanismen handeln.

Es ist wahrscheinlich, daß bei diesen verschiedenen Arten des Kritiktests dieselben Individuen sehr verschieden reagieren werden. Der visuelle Mensch wird optische Widersinnigkeiten schnell bemerken, über die ein abstrakt logisch gerichteter Mensch hinwegsieht usw. Hierzu kommt, daß die Kritikfähigkeit sehr stark vom Interesse abhängt; wer sich für technische Dinge stark interessiert, geht mit einer ganz anderen Einstellung an technische Dinge als an sprachlich-begriffliche heran und bemerkt daher auch die Fehler viel leichter.

Methodisch unterscheiden sich die Kritiktests dadurch, daß sie entweder einfache oder Wahlreaktionen verlangen.

Für niedere Entwicklungsstufen der Intelligenz muß man sich damit begnügen, einen einzelnen (sprachlichen oder optischen) Reiz vorzulegen und die Kritik herauszufordern durch die Fragen, ob sein Inhalt richtig oder falsch sei. Die oben gegebenen Beispiele von Sätzen und Bildern gehören hierher.

Höher steht die geforderte Leistung dort, wo die zu kritisierenden Elemente eingestreut sind in einen größeren sprachlichen oder bildlichen Zusammenhang; es wird dann nur die allgemeine Aufgabe gegeben, die zu bemängelnden Widersinnigkeiten herauszufinden. War im ersten Falle durch den Reiz und die Fragestellung schon die eindeutige Einstellung der Aufmerksamkeit auf die Absurdität gegeben, so muß im zweiten Falle bei jedem Glied des größeren Reizzusammenhanges erst die Entscheidungsfrage gelöst werden: Liegt hier eine kritikwürdige Stelle vor oder nicht?

Tests dieser Art sind (in der Form von Geschichten mit eingestreuten Widersinnigkeiten) in letzter Zeit insbesondere vom Hamburger Laboratorium (*8*, I u. II) ausgebildet und angewandt worden (MS. 131 bis 136). Sie scheinen sich für höhere Altersstufen und zur Prüfung höherer Begabungen sehr gut zu eignen; denn die geistige Leistungsfähigkeit tritt hier in recht verschiedenen Funktionen hervor, ohne daß die Vergleichbarkeit der Ergebnisse zu sehr beeinträchtigt wurde. Man kann feststellen: wie viele von den ausgesprochen sinnwidrigen Stellen des Textes gefunden werden, ob die Art ihrer Fehlerhaftigkeit erkannt und die Berichtigung in einsichtiger Weise vorgenommen wird, ob andere Stellen als sinnwidrig bezeichnet werden, die es gar nicht sind. Die Fassung der Texte bedarf noch mancher Verbesserungen; insbesondere wird darauf zu achten sein, daß durch die sprachliche Formulierung nicht unvermutete Schwierigkeiten bereitet werden, die die kritische Tätigkeit auf falsche Wege lenken können[1]).

Eine andere Form solcher Wahlkritiken schlägt neuerdings der Leipziger Lehrerverein (*197*) vor: Eine Reihe von Rechenaufgaben wird vorgelegt, unter denen sich einige unmögliche befinden; diese sollen herausgefunden werden (MS. 241—242).

Soweit sich bis jetzt vermuten läßt, gibt es bezüglich der Kritikfähigkeit drei typische Verhaltungsweisen der Prüflinge. Auf der einen Seite steht der Typus der „Stumpfen", welche die Sinnfehler nicht bemerken, alles Gegebene hinnehmen, als müßte es so sein. Diesem Typus der Kritiklosen steht nicht der eine Typus der kritischen Geister gegenüber, sondern zwei Typen: der Krittler und der Kritiker. Beide haben das Gemeinsame, daß sie die Neigung haben, Fehler zu bemerken und zu bemängeln. Beim Krittler aber entspricht dieser Neigung nicht eine gleiche Fähigkeit; er betätigt seine Neigung wahllos bei Wesentlichem und Unwesentlichem, bei wirklichen und bei nur vermeintlichen Fehlern, übersieht sogar oft wesentliche Mängel über gleichgültigen Kleinigkeiten, in die er sich verrennt. Der eigentliche kritische Geist aber beschränkt seine geistige Opposition auf solche Stellen, die es ver-

[1]) Andererseits muß man für höhere Altersstufen die Schwierigkeit noch erhöhen. Susanne Engelmann veröffentlicht soeben (*187*a) eine Stelle aus Paulsens Ethik, die durch Einfügung von 15 Absurditäten umgestaltet worden ist; der Test ist mit gutem Erfolg bei 19—20jährigen Lehramtskandidatinnen angewandt worden.

dienen, und läßt Unwesentliches beiseite. Sein intellektuelles Gewissen kann Sinnwidrigkeiten nicht vertragen; seine Unterscheidungsfähigkeit ermöglicht es ihm, nur auf wirklich gedanklichen Unsinn zu achten und nicht äußerliche Fehler (z. B. rein sprachlicher Art) damit zu verwechseln; seine Urteilsfähigkeit endlich setzt ihn in den Stand, den erkannten Sinnfehler auch richtig zu verbessern.

Diese Typen werden sich freilich erst in höherem Jugendalter allmählich mit einiger Deutlichkeit scheiden. Wie Versuche an 10jährigen recht begabten Kindern mit einem zusammenhängenden Text zeigen, steht diese Altersstufe im allgemeinen noch auf dem Niveau einer gewissen Stumpfheit in bezug auf rein gedankliche selbständige Kritikfähigkeit (*8*, I s. a. *218*, 8). Freilich konnten manche Kinder infolge des Zwanges der Aufgabe den Eindruck von „Krittlern" machen. Da sie kritisieren sollten und doch auf Grund ihres geistigen Entwicklungsstandes noch nicht korrekt zu kritisieren vermochten, griffen sie zu beliebigen und unwesentlichen Stellen, an denen sie der ihnen nicht recht angemessenen Aufgabe zu genügen suchten. Allerdings wurde auch hierdurch die überwiegende Stumpfheit nicht überdeckt; die Zahl der übersehenen kritikwürdigen Stellen ist meist überraschend groß.

14. Definitionen (MS. 138—144).

Die Definitionsversuche sollen nicht das Vorhandensein bestimmter Begriffe im kindlichen Bewußtsein feststellen — dazu dienen andere auf S. 88f. erwähnte Tests —, sondern sie sollen die Fähigkeit untersuchen, wieweit sich das Kind von den vorhandenen Begriffen Rechenschaft zu geben vermag, welche Merkmale des Begriffs von ihm vorwiegend beachtet werden und welcher Umfang dem Begriff zugeschrieben wird. Es dürfen also für diese Zwecke nicht Begriffe gewählt werden, deren Kenntnis zweifelhaft ist; denn wir wollen nicht die Kenntnisse des Kindes, weder die in der Schule noch die im Alltagsleben erworbenen, prüfen, sondern die Art und Korrektheit seines begrifflichen Denkens.

Die Aufgabe ist den Prüflingen zweifellos ungewohnt; denn wenn sie sonst in der Schule bestimmte Worte erklären sollen, so handelt es sich um solche, deren Bedeutung sie eben gelernt haben oder deren Gebrauch sie noch nicht beherrschen. Dagegen mutet sie die Frage nach Worten, deren Kenntnis sie durch den

Alltagsgebrauch längst erwiesen haben — wie Stuhl, Puppe, Onkel, Fenster, Neid, Mut usw. —, höchst sonderbar an; und es wird zuweilen nötig sein, durch eine vorauszuschickende Erklärung die Aufgabe einleuchtend zu machen (z. B.: sie sollten sich denken, ein Ausländer kenne den Sinn der Worte nicht und wolle ihn von ihnen erklärt haben).

Auch konkrete Beispiele wird man zweckmäßig vorausschicken, nie aber ein einziges allein, weil die hier vom Prüfer gewählte Definitionsform stark suggestiv auf die weiteren Lösungen des Prüflings wirken kann, — sondern mehrere Beispiele in verschiedener Fassung.

Feststellbar ist an den Leistungen einmal ihre psychologische Beschaffenheit, sodann ihr logischer Wert.

Entwicklungspsychologisch kommt hier vor allem die Scheidung der bloßen Zweckangabe von anderen Formen (Angabe des Materials, der Herkunft, des Gattungsbegriffs, eines Beispiels) in Betracht. Alle Untersuchungen zeigen übereinstimmend, daß die primitivste Form der Selbstbesinnung auf den Inhalt eines Begriffs in der Angabe des Zweckes besteht. Ein Stuhl ist „zum Sitzen", eine Puppe „zum Spielen", ein Soldat „zum Schießen", ja die Nase „zum Naseputzen". Nichts von begrifflichen Beziehungen der Unter- und Überordnung, nichts von der anschaulichen Beschaffenheit, nichts von der kausalen Bedingtheit des Gegenstandes wird auf dieser Stufe erwähnt; der Gegenstand ist vor allem da, um im praktischen Leben des Menschen einen bestimmten Dienst zu leisten. Zwischen dem 6. und 9. Jahre wird diese unvollkommene Stufe überwunden; nun wird versucht, zu dem gefragten Begriff den Oberbegriff zu finden (was ist ein Stuhl? ein Möbel), oder sein Material anzugeben (was ist ein Schlüssel? aus Eisen), oder ein Beispiel zu nennen (was ist Beute? wenn einer Geschütze erobert hat).

Ein Altersfortschritt zeigt sich ferner in der stufenweisen Bewältigung verschiedener Begriffsarten. Hierfür liegen u. a. Massenuntersuchungen von Gregor (*65*) vor, deren Ergebnisse von mir (*114*) nach Altersstufen umgerechnet wurden (MS. 141 bis 142). Von kleineren Kindern bis etwa zum 11. Lebensjahr kann man lediglich die Definition konkreter Gegenstände verlangen; abstrakte Begriffe werden kaum von ihnen verwandt, geschweige denn so beherrscht, daß sie darüber Rechenschaft geben könnten. Vom 6. Schuljahr an sind die Kinder imstande, halb-

abstrakte Begriffe — z. B. solche, die politische, soziale, rechtliche Verhältnisse ausdrücken, wie Gesetz, Bündnis, Ordnung, Obrigkeit, Geld — zu erklären. Ganz abstrakte Begriffe logischer oder ethischer Natur (z. B. Urteil, Erklärung, Gerechtigkeit, Mitleid usw.) werden erst im 7. und 8. Schuljahr (13. und 14. Lebensjahr) zureichend definiert, und auch dies gilt nur von den leichteren Begriffen dieser Art; andere sind noch weit über das Schulalter hinaus zu schwer. — Wegen der Schwierigkeit der Definition halbabstrakter und abstrakter Begriffe sind diese für die Untersuchung der höheren Begabung besonders geeignet. Aber auch unter den konkreten Ausdrücken gibt es manche recht schwere (z. B. Fenster); auch Verwandtschaftsbezeichnungen (Onkel, Vetter) werden nicht ganz leicht in ihren Wesensmerkmalen getroffen.

In der Art der Beantwortung kann man einen mehr speziellen und einen mehr allgemeinen Typ unterscheiden. Es sondern sich also die Konkretdenker von den Abstraktdenkern, ohne daß wir aber hieraus eine Wertabstufung ableiten dürften. Das Denken der ersten arbeitet mit lebhaften Anschauungsbildern, die freilich die Weite des Begriffsumfangs nicht erkennen lassen, das der anderen mit begrifflichen Allgemeinheiten, denen die anschauliche Verlebendigung fehlt und die zuweilen nichts als Worte ohne scharf gedachte Bedeutung sind. Ja, unter Umständen kann eine ganz allgemeine Definition geradezu eine Minderwertigkeit gegenüber der spezielleren bedeuten; auf die Frage: „Was ist ein Briefträger?“ ist die Antwort: „Ein Mensch“ viel bequemer und leerer als die Antwort: „Ein Beamter“.

Roloff, der an etwa 1000 Kindern Definitionsversuche gemacht hat (MS. 142), gibt eine Reihe von lehrreichen Proben für die beiden Typen. So lautet für die Definition von Mut eine allgemeine Lösung: „Dann hat man keine Angst;“ eine spezielle: „Wenn einer auf Patrouille geht und wagt sich weit vor.“ Für Beute eine allgemeine Lösung: „Wenn unsere Soldaten etwas erobert haben;“ eine spezielle: „Wenn der Adler sich einen Vogel fängt.“

Die Statistik ergab, daß die Prozentzahlen der allgemeinen Lösungen für die Begriffe Mut, Neid und Beute vom 9. bis zum 12. Jahre keinen merklichen Zuwachs zeigen; also auch nach diesem Befund ist die allgemeinere Definition kein Kennzeichen fortschreitender Entwicklung, wenigstens nicht während jener Altersspanne. Das wird erst in und nach der Pubertät anders, wenn eigentlich logische Motive auftreten und eine möglichst exakte mit Oberbegriffen und differentia specifica arbeitende Definition versucht wird.

Will man die Definitionsleistungen zahlenmäßig bewerten, so wird man ihren logischen Gehalt zugrunde legen müssen. Dabei kommt es allerdings durchaus nicht darauf an, ob die Definition den Regeln der Schullogik entspricht — gerade die kindlichen Leistungen zeigen ja, wie lebensfern und künstlich jene Regeln sind —, sondern nur darauf, ob die Abgrenzung des Begriffs und die hervorgehobenen Merkmale eine Annäherung an die logische Norm des Begriffs zeigen. In dieser Annäherung gibt es alle möglichen Stufen: der Umfang des Begriffs kann mehr oder weniger verwaschen oder verschoben werden; als Merkmale können wesentliche oder unwesentliche genannt werden; die Ausdrucksweise kann scharf oder unklar sein. Man wird hiernach etwa folgende Wertskala der Definitionen aufstellen können: unsinnig, fehlerhaft, unklar, angängig, richtig, logisch vollkommen; doch kann man sich für viele Zwecke mit einer einfacheren Bewertung, etwa den drei Stufen „falsch, angängig, richtig" begnügen.

Neuerdings hat Gregor (*65a*) versucht, den Definitionstest so zu eichen, daß mit ihm allein das „Intelligenzalter" des geprüften Kindes bestimmt werden kann. Er arbeitet mit 37 Begriffen; die gelieferten Definitionen teilt er in „korrekte", „richtige", „primitive" und „falsche" ein. Eine Durchprüfung an verschiedenen Schulklassen ermöglichte es ihm, für jeden Begriff und jede Klasse die „charakteristische" (d. h. von mehr als der Hälfte der Kinder angewandte) Lösungsform festzustellen. Wird nun ein Kind mit allen 37 Begriffen geprüft, so weisen die Lösungsformen auf bestimmte Klassenstufen hin, aus deren Gesamtheit dann das „I.-Alter" abgeleitet wird.

Es scheint nun gewiß wertvoll, jedem mit dem Definitionstest geprüften Kind einen Leistungsindex geben zu können; aber grundsätzlich muß bestritten werden, daß dieser Index die Bedeutung des „I.-Alters" habe. Denn die Höhe der Intelligenz als einer Gesamtfähigkeit kann niemals durch einen Einzeltest, und sei er noch so vorzüglich, sondern nur durch eine Kombination sehr verschiedenartiger Tests gemessen werden. (Vgl. die Ausführungen über das I.-Alter in Kap. VII, 3.)

Aus den noch unveröffentlichten Massenuntersuchungen, die Roloff mit dem Definitionstest an Bergedorfer Schulkindern anstellte, seien die folgenden Ergebnisse schon jetzt angedeutet[1]): Wurden für die Schüler von 12 Volksschulklassen die Durchschnittsleistungen von je 14 Definitionen berechnet und die Schüler dann nach der I.-Schätzung der Lehrer in eine Rangordnung gebracht, dann hatte das beste Viertel einen Vorsprung von 1½ Intelligenzjahren vor dem schwächsten Viertel. Die Korrelation des Tests zur geschätzten I. ist also bedeutend. — Vergleicht man ferner die Altersstufen von 9—14 Jahren, so ist ein gleichmäßiger stetiger Anstieg der De-

[1]) Die Veröffentlichung wird voraussichtlich in den „Hamburger Arbeiten zur Begabungsforschung" erfolgen.

finitionsleistungen erkennbar. — Die ebenfalls deutlichen Korrelationen der Testergebnisse zu den sozialen Unterschieden der Kinder werden im Kap. erwähnt.

15. Schließen (MS. 144—150).

Auch hier ist ein aus der Logik stammendes Verfahren auf die Psychologie übertragen worden. Aber noch weit mehr als bei den Definitionen muß man sich beim Schließen davor hüten, die Kunstgebilde der Schullogik als Kennzeichen und Maßstäbe für den natürlichen Denkvorgang und seine Leistungsfähigkeit anzusehen.

„Schlüsse", d. h. Ableitungen von Urteilen aus gegebenen Voraussetzungen, vermag das Kind schon sehr früh zu vollziehen; in der „Psychologie der frühen Kindheit" (*13*) bespreche ich (S. 270f.) eine ganze Reihe von Denkvorgängen, die zweifellos den Charakter von Schlüssen haben. Und auch in vielen der bisher beschriebenen Tests zur Prüfung der theoretischen wie der praktischen I. werden fortwährend Schlüsse sehr verschiedener Art vollzogen. Freilich verlaufen sie durchweg in anderer Form, als die aristotelische Logik sie uns in den vier Schlußfiguren vorführt. Diese Figuren sind nicht der Ausdruck dafür, wie sich das Fortschreiten unseres Denkens tatsächlich vollzieht, sondern Maßstäbe, nach denen wir die Geltung von Ergebnissen (ganz gleich auf welchem Wege diese gefunden sein mögen) nachträglich zu rechtfertigen imstande sind. Nichts wäre verkehrter als hier logische Norm mit psychologischer Tatsächlichkeit zu verwechseln.

Nun hatte Meumann (*9*, Bd. I, S. 548) seinerzeit die Behauptung aufgestellt, daß „die eigentlich logische Schlußfolgerung, ausgeführt in der Form, wie sie sich im Schulschluß vollzieht," von dem Kinde nicht vor dem 14. Lebensjahr vollzogen wird. Ich hatte dieser Auffassung auf dem VI. Psychologenkongreß entschieden widersprochen. Daraufhin hat Schüßler (*109*) eine experimentelle Nachprüfung vorgenommen; er untersuchte seine Schülerinnen dreimal in Jahresabständen: als 12-, 13- und 14jährige, und zwar mit Schlüssen aller vier Figuren, bei denen die Prüflinge jedesmal zu den gegebenen Prämissen die Schlußfolgerungen zu finden hatten (MS. 145—148). Ich gebe zu jeder Schlußfigur je ein Beispiel:

I. Die Säugetiere haben warmes Blut.
Der Walfisch ist ein Säugetier . . .

II. Kein See ist ein fließendes Gewässer.
Jeder Bach ist ein fließendes Gewässer . . .
III. Die Adler fressen Fleisch.
Die Adler sind Vögel . . .
IV. Die Löwen sind Katzen.
Die Katzen sind Raubtiere . . .

Es ergab sich, daß die Schlüsse der ersten Figur die leichtesten waren; 60% bis 90% der Kinder vermochten (wenn auch nicht in allen Fällen) Schlüsse dieser Art zu ziehen; die entsprechenden Ziffern liegen bei der zweiten Figur um 35%, bei der dritten um 20% herum, während bei der vierten nur ganz ausnahmsweise Treffer vorkamen.

Dies starke Versagen beruht nun nach meiner Meinung vor allem auf der Unnatürlichkeit der Denkaufgaben; die erste Schlußfigur ist zweifellos diejenige, die mit dem natürlichen Denkablauf die größte Ähnlichkeit hat; daher lieferte sie die besten Ergebnisse. Daß das Versagen der Kinder nicht, wie Meumann glaubt, auf ihrer noch unentwickelten Denkfähigkeit beruht, konnte Schüßler auf meine Anregung durch Parallelversuche an Erwachsenen aus dem Volke erweisen; diese leisteten nicht nur nichts Besseres, sondern zum Teil Geringeres als die Kinder — ein Zeichen, daß das Denken in den Bahnen der Schlußfiguren überhaupt der natürlichen Denkweise widerstrebt.

Die Lösbarkeit eines Schlusses hängt übrigens nicht nur von der Schlußfigur ab, sondern auch von der Art der beteiligten Begriffe; sind diese dem Kinde geläufig, so vollzieht sich der Schluß leichter.

Für IP. kommen nach Schüßler höchstens Schlüsse nach der ersten Figur in Betracht, die mit klaren und bekannten Begriffen arbeiten; hier zeigt sich der Hauptaltersfortschritt vom 11. zum 12. Jahr (während vom 13. zum 14. ein Stillstand der Leistungen zu verzeichnen war).

Nicht ausgeschlossen erscheint es, daß für höhere Alters- und Begabungsstufen Schlüsse von abstrakterer und komplizierterer Form brauchbar sind. Unsere Methodensammlung bringt daher einige Beispiele von Schlüssen, die von Störring (*116*) und Lindworsky (*80*) zu denkpsychologischen Untersuchungen an Erwachsenen angewandt worden sind (MS. 148—150); ihre Verwendbarkeit für Fähigkeitsprüfungen ist erst zu erproben. Wertvoll scheinen insbesondere Vexierprämissen zu sein, d. h. solche, die in Wirklichkeit keinen Schluß erlauben, wie etwa:

U ist links von V.
F ist links von V.

Ein solcher Test würde der Vorlage von unlösbaren Rechenaufgaben ähneln, die wir S. 103 erwähnten.

Ferner wäre es interessant festzustellen, ob, bzw. wann die Prüflinge zwischen der bloßen Tatsächlichkeit des Schlußergebnisses und seiner logischen Notwendigkeit zu unterscheiden wissen. Doch sehe ich noch keine geeignete Methode für diese Feststellung.

Als Nachtrag sei ein Hinweis auf eine soeben erschienene Mitteilung Deuchlers (*59a*) angefügt, die die Prüfung des kindlichen Schließens zu fördern geeignet ist. Nach allgemeinpsychologischen Erörterungen über das schließende Denken schildert D. das von ihm mit Erfolg angewandte Prüfverfahren und zählt 38 geeignete Schlußaufgaben auf. Ergebnisse werden noch nicht mitgeteilt.

16. Ergänzungstests (MS. 150—170).

Der Test der Lückenergänzung hat eine geschichtlich bedeutsame Rolle in der Entwicklung der IP. gespielt; war er doch die erste Prüfungsmethode, die ausdrücklich zur Untersuchung der höheren geistigen Leistungsfähigkeit vorgeschlagen wurde. Als Ebbinghaus (*60*) die Methode 1897 erstmalig anwandte (MS. 151), setzte er freilich zu weitgehende Hoffnungen in sie; er war der Meinung, das Wesen der geistigen Leistungsfähigkeit sei nichts anderes als Kombinationsfähigkeit, und diese könne dadurch festgestellt werden, daß man die in einem zusammenhängenden Text gelassenen Lücken sinnvoll ergänzen lasse. Wir wissen heute, daß mit dem Test nie die Intelligenz, sondern nur ein eng begrenztes Teilgebiet derselben geprüft wird (s. S. 49ff.); in diesen Grenzen aber ist der Test sehr wertvoll; auch hat sich gezeigt, daß seine Ergebnisse mit dem Lehrerurteil über die I. der Schüler in hoher Korrelation stehen.

Geprüft wird diejenige Seite der I., die mit der sprachlichen Befähigung in enger Verbindung steht. Denn sowohl das Verständnis eines nicht ganz vollständigen Textes wie auch das Finden der geeigneten Lückenausfüllung ist zum Teil von der Feinheit des Sprachgefühls und von Reichtum und Bereitschaft des Wortschatzes abhängig. Dies ist kein Schade, sobald man sich von dieser Komplikation die nötige Rechenschaft gibt; denn die Fähigkeit, einen gefundenen Gedanken durch sprachliche Formulierung möglichst scharf und eindeutig zu bestimmen, gehört auch zur I. Allerdings muß man durch richtige Gestaltung des Lückentextes dafür sorgen, daß bei der Ausfüllung wirkliche Denkarbeit geleistet wird. Ist die Ausfüllung zu leicht gemacht, dann wird die Lücke rein mecha-

nisch ergänzt; d. h. es laufen die durch sprachliche Übung ausgeschliffenen Assoziationen ab, ohne daß ein eigentliches Nachdenken nötig wäre. Ist andererseits Zahl und Umfang der Lücken zu groß, dann handelt es sich nicht mehr so sehr um ein durch feste Ziele bestimmtes Ergänzen, als um ein vages phantasiemäßiges Kombinieren; und die Aufgabe nähert sich zu sehr einem Rätselspiel.

Die älteren Ergänzungsmethoden (MS. 150—155 oben) haben diese beiden Mängel nicht immer vermieden, wie sie überhaupt in der Durchlöcherung des Textes allzu unsystematisch und willkürlich vorgingen. Wir finden da nebeneinander in demselben Text ausgelassene Buchstaben, Silben und ganze Wörter, und unter diesen wiederum Wörter verschiedenster Art. Hierdurch war nicht nur die Denkschwierigkeit, sondern auch die ganze psychologische Beschaffenheit des Ergänzungsvorgangs von Lücke zu Lücke so verschiedenartig, daß eine wirkliche Vergleichung der einzelnen Ausfüllungen und damit eine exakte Bewertung der Gesamtleistung eines Prüflings nicht einwandfrei möglich war.

Als methodischer Fortschritt hatte es daher zu gelten, daß man nur Worte einer bestimmten Kategorie fortließ; hierdurch wurde eine gemeinsame Denkeinstellung für den ganzen Test gewährleistet, ohne daß doch eine Mechanisierung möglich war; denn jede Lücke stellte eine individuelle Aufgabe dar, die für sich überdacht werden mußte.

Die leichtere Form dieses neuen Verfahrens arbeitet mit der Ergänzung von Verben (MS. 157—161); es muß also innerhalb der Gedankeneinheit des einzelnen Satzes die tragende Handlung richtig erkannt werden (Mayer (*84*), Winteler (*124*), Hylla). Erschwert wurde der Test zum Teil dadurch, daß man nicht einmal die Stelle der Auslassung kennzeichnete, sondern die verstümmelten Sätze ohne sichtbare Lücken vorlegte.

Eine Abwandlung dieser Methode besteht darin, daß man bei jeder Lücke nicht ein Verb, sondern eine möglichst große Zahl synonymer Verben verlangt (MS. 158 unten und 160—161). Z. B. in dem Satz: „Die Fußgänger —, um schnell nach Hause zu kommen" ist zu ergänzen: „liefen, eilten, rannten". Hierbei wird ja in erster Linie der Reichtum des Wortschatzes geprüft; nach den Erfahrungen Wintelers und unseres Hamburger Laboratoriums scheint aber das Finden zutreffender Synonyma und das Meiden ungeeigneter Ersatzwörter in einer engen Korrelation zur I. zu stehen.

Viel schwerer ist die Ergänzung von Konjunktionen (MS. 161—170); dies zuerst von Lipmann (*82*) angewandte, dann von Minkus (*31*) in Breslau ausgebaute Verfahren hat sich zur Prüfung von höheren Alters- und Begabungsstufen vorzüglich bewährt. Die Bindewörter (da, damit, obgleich, daher, darauf, trotzdem usw.) geben das logische Verhältnis zweier Gedankeneinheiten zueinander an; diese Beziehung der Finalität oder der Begründung, des Gegensatzes oder der Zeitfolge usw. richtig zu verstehen und auszudrücken, erfordert eine ausgesprochene Denkleistung höherer Art. In meiner Bearbeitung der Minkusschen Ergebnisse[1]) konnte festgestellt werden, daß der Test insbesondere für die Jahrgänge 12—14 einen starken Altersfortschritt, zugleich eine noch stärkere Differenzierung der gut und schlecht befähigten Schüler gleicher Jahrgänge aufzeigt; bei höheren Jahrgängen begannen auch die äußeren Lebensbedingungen (fremdsprachliche Schulung, Vorbereitung zu höheren Berufsformen) eine Rolle zu spielen. Es ist daher angebracht, mit dem Test nur solche Prüflinge untereinander zu vergleichen, die in bezug auf ihre sprachliche Kultur als einigermaßen gleichwertig gelten können. Außerdem ergab sich, daß die verschiedenen Kategorien von Bindewörtern verschiedene hohe Denkansprüche stellten: die temporalen und kausalen Gedankenverknüpfungen boten einen geringeren Grad psychologischer Schwierigkeit als alle übrigen; die am schwersten zu ergänzenden Bindewörter waren die des Gegensatzes (obgleich, statt dessen usw.).

Die Verbindung einer Ergänzung und einer Auswahltätigkeit verlangt die „Textergänzung nach Wortlisten", die von Minkus (MS. 168—169) und in veränderter Form vom Hamburger Laboratorium angewandt wurde (MS. 155—156). Im letzten Falle wird zu jeder Lücke eine Liste von vier bis sechs Worten vorgelegt, unter denen sich das richtige findet. Dies Heraussuchen hat eine gewisse Ähnlichkeit mit dem Vokabelaufschlagen bei Übersetzungen aus einer fremden Sprache; denn auch hier handelt es sich darum, aus den vielen Bedeutungen, die bei einer unbekannten Vokabel im Lexikon stehen, diejenige herauszufinden, die sich dem Zusammenhang des Textes einpaßt. Der noch nicht ganz verstandene Zusammenhang und die zu suchende Wortbedeutung müssen zusammenschießen zu einem nun voll verständlichen sinnvollen Ganzen.

[1]) Vgl. dieses Buch S. 30f.

Der Umstand, daß wir an mehreren Stellen der Besprechung auf den Anteil der sprachlichen Fähigkeiten bei Lösung der Ergänzungstests hinweisen mußten, legt den Gedanken nahe, diese Testform zu einem speziellen Untersuchungsmittel der sprachlichen Begabung auszubauen. Solche Tests werden ja notwendig werden, wenn es sich um die Auslese für den Sprachenzug der höheren Schulen handelt.

17. Dreiwort- und Zweiwortmethode (MS. 170—175).

Der Test „Bildung eines Satzes aus mehreren Stichwörtern" ist von Masselon in die Psychiatrie, von Binet (*3*, *4*) in die Psychologie eingeführt worden; in Deutschland haben Piorkowski (*90*), Meumann (*86*) u. a. den Test ausgebaut und seine Ergebnisse näher bestimmt.

Vorgelegt werden bald zwei Worte, bald drei Worte; neuerdings hält man sich nicht streng an einzelne „Worte", sondern bietet Gedankeneinheiten, deren jede aus einem oder mehreren Worten bestehen kann, z. B.: „Fast verdursteter Hund, mitleidiger Mensch, verhinderter Überfall" (Piorkowski). Es empfiehlt sich in solchen Fällen, die Prüflinge darauf hinzuweisen, daß sie sich an die grammatische Form der Darbietung nicht buchstäblich zu halten brauchen.

Ebenso ist auch der Ausdruck „einen Satz zu bilden" nicht in seiner grammatischen Strenge zu nehmen. Es soll ein einheitlicher gedanklicher Zusammenhang geschaffen werden; ob dieser in einem oder mehreren Sätzen ausgedrückt wird, ist von geringerer Bedeutung. Meumann verlangte deshalb, daß aus den Stichworten „eine kleine Geschichte" gebildet werde; er bot allerdings hierfür eine größere Anzahl von Stichworten.

Der Masselon-Test hat mit dem im vorigen Abschnitt beschriebenen Ergänzungstest eine gewisse Ähnlichkeit; aber während dort der Text doch in der Hauptsache gegeben war und nur die vereinzelten Lücken Anlaß zur Denkarbeit boten, sind hier nur wenige Textinseln gegeben, und der ganze Satzsinn und Wortlaut muß vom Prüfling selbst gestaltet werden. Infolgedessen kann sich hier die konstruktive Phantasie, die im vorigen Test ausgeschaltet sein sollte, betätigen; aber der Anteil dieses mehr phantasiemäßigen Verhaltens an der Leistung ist sehr verschieden. Einige Forscher glauben daher, hier geradezu zwei qualitativ verschiedene

Typen feststellen zu können: einen mehr logischen und einen mehr phantasievollen. Piorkowski schildert sie (S. 60) folgendermaßen:

„Die einen verbinden ihre Vorstellungen streng sachlich mit logischer Einfachheit; dabei haben die Verbindungen etwas Konventionelles an sich, sind selten neu, niemals sprunghaft. Ihnen stehen andere gegenüber, deren Eigenart gerade in den originellen, aufs erste oft verblüffenden Kombinationen besteht. Sie bevorzugen sprunghaftes Vorwärtsbewegen der Gedanken, benutzen zu neuen Verbindungen Hilfsglieder, die manchmal fast als fremde Elemente zu betrachten sind. Auf der einen Seite der Denker, auf der anderen der Phantast im guten Sinne."

Es ist klar, daß bei solchen qualitativen Unterschieden die quantitative Bewertung und damit die Benutzung des Tests zur Feststellung von I.-Graden sehr erschwert wird. Ob man die phantastisch originelle und wortreiche, oder ob man die einfache und knappe, aber logisch korrekte Leistung höher einschätzen solle, hängt stark von subjektiven Neigungen des Prüfers oder von den besonderen Zielen der Untersuchung ab.

Piorkowski und das Hamburger Laboratorium (MS. 172 bis 175) führten eine Skala von vier bzw. fünf Stufen in der Bewertung der einzelnen Lösungen ein.

Die Schwierigkeit der kombinatorischen Leistung hängt natürlich in hohem Grad von den gewählten Stichworten ab. Es gibt solche, die den zu bildenden Satz fast eindeutig bestimmen und nur ganz geringe Denkarbeit erfordern (z. B.: Jäger, Wald, Hase), andere, bei denen die Worte zunächst gleichgültig zueinander sind oder gar in Widerspruch miteinander zu stehen scheinen. Im letzteren Fall (Beispiel: stehengebliebene Uhr, geschehenes Eisenbahnunglück, Freude) muß erst eine Hemmung überwunden werden, um die beiden kontrastierenden Begriffe Unglück und Freude in Zusammenhang miteinander zu bringen. Diese nicht unbeträchtliche Denkleistung kann, wie wir in Hamburg feststellten, gelegentlich schon von begabten 10jährigen vollbracht werden. (Die stehengebliebene Uhr führt zum Versäumen eines Zuges, dem dann das Unglück zustößt; da ist die Freude groß, nicht dabei gewesen zu sein.)

Sehr lehrreich sind diejenigen Stichwortgruppen, bei denen ein weiterer wesentlicher Begriff hinzugedacht werden muß, damit eine wirklich pointierte Lösung zustande komme. Hierauf wies schon Meumann hin: Die Zweiwortgruppe „Esel—Schläge" kann zu dem pointenlosen Satz „Der Esel bekommt Schläge" und zu dem pointierten „Der Esel bekommt wegen seiner Faulheit

Schläge" führen. Die zweite Lösungsweise ist, wie mir scheint, nicht nur ein Zeichen höherer Intelligenz, sondern auch stärkerer Intellektualität; das Kind begnügt sich nicht mit der bloß zeitlichen Aneinanderreihung beider Begriffe, sondern hat das Bedürfnis nach ihrer kausalen Verknüpfung. Auf höherer Stufe beobachteten wir in Hamburg ähnliches bei dem Test: „Soldaten im Lager — sternlose Nacht — große Verwirrung." Unter begabten 10jährigen fanden sich immerhin 25%, die erwähnten, wie die sternlose Nacht die Verwirrung im Lager verursachen konnte (indem sie etwa einen Überfall durch den Feind oder das unbemerkte Nahen eines Fliegers ermöglichte).

Eine neue Wendung erhielt der Test durch Moede und Piorkowski (*206*), indem für jede Stichwortgruppe mehrere abweichende Lösungen verlangt wurden (MS. 175). Die Erschwerung ist beträchtlich, daher ist das Verfahren nur für höhere Alters- und Begabungsstufen brauchbar. Jetzt kommt es nicht mehr in erster Linie auf die Güte der einzelnen Lösung, sondern auf die Verschiedenartigkeit der Lösungen untereinander an; es wird also die Beweglichkeit des Vorstellungslebens geprüft. Manche Prüflinge vermochten von einer einmal gefundenen Gedankenverknüpfung nicht loszukommen und brachten nur unbedeutende Abwandlungen der ersten Lösung zustande; andere zeigten eine bemerkenswerte Mannigfaltigkeit der Lösungen.

Mehrfache Lösungen verlangte auch Penkert (*8*, II), aber in einer anderen Form (MS. 207). Er ließ aus einer größeren Stichwortreihe eine Geschichte bilden und niederschreiben, und zwar einmal in kurzer, schlichter Sachlichkeit, sodann in lebendig anschaulicher Darstellung. Das Verfahren scheint geeignet, den oben besprochenen Typenunterschied zwischen dem mehr logischen und dem mehr phantasievollen Verhalten besonders deutlich herauszustellen, da man bald erkennen wird, welche Art der Beschreibung dem Prüfling besser liegt.

Der Test gehört zweifellos zu den besten, über welche wir verfügen. Er zeigt bedeutende Altersfortschritte, insbesondere in der Zeit vom 9. bis zum 12. Lebensjahr, in den schwereren Formen auch bei höheren Altersstufen; zugleich aber treten innerhalb gleichaltriger Prüflinge ganz gewaltige Begabungsdifferenzen zutage.

Die praktische Verwertung des Tests wird dadurch gefördert, daß seine Schwierigkeit in weitem Umfange abstufbar ist, so daß er den verschiedensten Aufgaben und Altersstufen angepaßt werden kann.

18. Finden des Wesentlichen (MS. 175—182).

Die Analyse, die das Denken an irgend einem zusammengesetzten Inhalt vernimmt, muß zugleich stets eine Auslese sein. Zahlreiche Elemente sollen ausgeschieden werden, damit die für den vorliegenden Denkzweck bedeutsamen Inhalte, die sogenannten „wesentlichen" Bestandteile, um so schärfer hervortreten. Von dieser Auslese war bereits dort die Rede, wo wir die Merkfähigkeit für Geschichten und Bilder behandelten (S. 77 u. 94). Sie tritt aber noch viel deutlicher zutage, wenn die Aufgabe selbst eine möglichste Zusammendrängung des Stoffes erfordert. Wenn es gilt, die Hauptgedanken eines Textes in einen oder wenige knappe Sätze zu fassen, oder eine recht treffende Überschrift für eine Geschichte, eine Unterschrift unter ein Bild zu finden, so muß sich in der Beschränkung der Meister zeigen; „wenig aber wesentlich" heißt die Forderung, die nur ein Mensch mit wirklich starker Denkfähigkeit erfüllen kann.

Für diese Verdichtungsleistung hat soeben der Leipziger Lehrerverein (*197*) einige Tests vorgeschlagen (MS. 239—241). Der eine besteht in der Darbietung einer langen Geschichte, deren Hauptpunkte durch zahlreiche vom Wesentlichen ablenkende Ausschmückungen umkleidet sind. Die Anweisung lautet dann: „das Wichtigste in einigen Sätzen sofort niederzuschreiben". Ähnlich wird mit einem naturwissenschaftlichen Stoff verfahren, nur daß hier, um das Gedächtnis nicht zu stark mitwirken zu lassen, jeder Abschnitt für sich zu einem kurzen Satz verdichtet werden muß. Besonders aussichtsvoll erscheint mir der „Telegrammtest": ein längerer Brief soll in ein Telegramm verwandelt werden. Hier ist der Zwang zur Kürze und der völligen Ausschaltung alles Unwesentlichen dem Prüfling besonders leicht plausibel zu machen.

Eine weitere Verwicklung erfährt die Denkaufgabe, wenn nicht nur Verdichtung, sondern zugleich auch Verallgemeinerung des Stoffes gefordert wird. Sehr oft nämlich ist der Kerngedanke eines Inhalts viel umfassender, als dessen unmittelbar konkrete Fassung selbst erkennen läßt; die wahre Pointe ist unter einem Gleichnis, einer Umschreibung, einem Einzelbeispiel verborgen und muß aus diesen Verhüllungen herausgeschält werden; oder sie ist selbst gar nicht ausgedrückt, sondern soll auf Grund der Vorbereitungen des Textes selbst gefunden werden.

Hierher würde die Aufgabe gehören, die Pointe eines Witzes zu erfassen; aber das Witzverständnis ist vermutlich eine ver-

wickeltere Fähigkeit, an der nicht nur die algemeine I.-Höhe, sondern auch eine spezifische, bisher noch nicht näher untersuchte Begabung für Humor, Paradoxie und Überraschung gehört. Daher kommt es, daß viele zweifellos intelligente Menschen Witzen gegenüber merkwürdig hilflos sind[1]).

Für die Zwecke der IP. ist es daher besser, den Paradoxiefaktor auszuschalten und bei einem ernsthaften Stoff die Herausarbeitung des verallgemeinerten Grundgedankens zu verlangen. Die geeignetsten Hilfsmittel hierfür sind Sprichwörter und Fabeln. Die ersteren (MS. 181—182) sind für Prüfungen von Kindern und Jugendlichen noch kaum verwandt worden; will man es tun, so empfiehlt es sich, unbekannte Sprichwörter (etwa die eines anderen Volkes) zu nehmen, da bei bekannten leicht schon unkontrollierbare Unterhaltungen über den Inhalt vorausgegangen sein könnten.

Fabeln (MS. 175—181) sind von Terman und Childs (*118*) und neuerdings im Hamburger Laboratorium (*8*, I u. II) mit gutem Erfolg angewandt worden. Sie haben den Vorteil, daß sie ihrer ganzen Anlage nach darauf berechnet sind, in eingekleideter Form „die Moral von der Geschicht'", also einen allgemeinen ethischen Gedanken dem Leser nahezubringen. Wann sind Kinder zu dieser gedanklichen Leistung fähig?

Die Versuche zeigen, daß die schwächeren Intelligenzen von der geforderten Lösung nach zwei Richtungen abweichen. Entweder bleiben sie ganz am konkreten Inhalt haften und geben statt der „Lehre" einen kurzen Auszug der Fabel; oder aber sie bringen eine Nutzanwendung von ganz vager Allgemeinheit, die an dem spezifischen Moralgedanken der Fabel vorbeigeht. Jedenfalls ist bei Fabeln deutlicher als bei irgend welchen anderen Stoffen zu erkennen, ob das Wesentliche des Gedankens erfaßt worden ist oder nicht.

Unintelligente Kinder zeigen ferner eine überraschende Unselbständigkeit in der Formulierung der „Lehre". Als Vorbereitung war bei unseren Hamburger Versuchen die Fabel „Der Rabe und der Fuchs" mitsamt der Erklärung dargeboten worden, welche lautete: „Glaube nie dem Schmeichler; denn er will nur etwas von dir haben". Diese Fassung führte suggestiv manche Kinder dazu, bei den nun folgenden zwei ganz anders gearteten Fabeln die Lehre so zu formulieren: „Glaube nicht dem Rübezahl, denn er will dir was antun." „Glaube nur dem Helfer, er gibt dir was wieder."

[1]) Eine ähnliche Kombination allgemeiner I. mit einer noch wenig bekannten Spezialfähigkeit liegt auch bei der Begabung zum Rätselraten vor.

19. Vergleichen (MS. 182—194).

Gleiches und Verschiedenes kann im Bewußtsein Wirkungen erzeugen, ohne daß ein Vergleichen und Unterscheiden vorliegen müßte. So kann der Anblick eines Bildes das Kind veranlassen, es „Mama" zu benennen, lediglich auf Grund einer Ähnlichkeitsassoziation. Das wirkliche Vergleichen (bzw. Unterscheiden) ist eine ausgesprochen intellektuelle Tätigkeit; sie setzt voraus, daß die gedanklichen Kategorien „gleich", „ähnlich", „dasselbe" oder „anders", „unähnlich", „verschieden" an zwei Bewußtseinsinhalte angelegt werden. Die einfachere Form ist das „zweigliedrige" Vergleichen, das die beiden Vergleichsobjekte in ihrer komplexen Ganzheit als gleich, ähnlich oder verschieden auffaßt, ohne sich Rechenschaft darüber zu geben, in welchen Merkmalen sie übereinstimmen oder abweichen. Weit höher steht das „dreigliedrige" Vergleichen, welches ein „tertium comparationis" anwendet; hier wird außer den beiden Gegenständen noch das Merkmal, in bezug auf welches verglichen wird, ins Bewußtsein gehoben. Die Isolierung des Vergleichsmerkmals ist nur möglich durch Absehung von den übrigen für diesen Zweck gleichgültigen Merkmalen, also durch einen Akt der Abstraktion. Die höheren Vergleichsleistungen sind deshalb zugleich auch Abstraktionsleistungen.

Die Schwierigkeit des Vergleichens ist von zwei Bedingungen abhängig. Die eine, der Grad des Unterschieds, hat mit unserem Problem nicht unmittelbar zu tun; ob noch ein geringerer Grad eines Helligkeits- oder Größe- oder Schwereunterschieds gemerkt wird oder nicht, hängt von der Unterschiedsempfindlichkeit, der Aufmerksamkeit, dem Interesse, aber nicht von der I. ab. Um so mehr gehört die andere Bedingung: die Isolierbarkeit des Vergleichsmerkmals, hierher.

Am leichtesten sind solche Vergleichungen zu vollziehen, bei denen das Vergleichsmerkmal in einem selbständig wahrnehmbaren Bruchstück der Vergleichsgegenstände besteht. Hiervon war bereits bei den Heilbronner-Serien die Rede (s. S. 95): schon sehr kleine Kinder sind imstande, in zwei Umrißzeichnungen desselben Gegenstandes den einzelnen Strich herauszufinden, durch den die eine vollständiger ist als die andere. Schwerer ist die entsprechende Aufgabe in den Vergleichtests von Koch und Habrich (*74*, *66*): zwei Gruppen geometrischer Figuren werden gezeigt; einige Figuren sind beiden Gruppen gemeinsam, andere sind verschieden. Aus einer größeren Zahl von Figuren müssen dann die „gleichen"

und die „verschiedenen" Figuren der beiden Gruppen herausgefunden werden (MS. 184—185). Der Test (der von den Veranstaltern geradezu als Mittel zur Prüfung der „Abstraktionsfähigkeit" hingestellt wird) zeigte eine deutliche Korrelation sowohl mit dem Alter wie mit der Begabung der Kinder.

Auf einer anderen Stufe stehen Vergleiche, bei denen formale oder materiale Eigenschaften der Gegenstände verglichen werden sollen — sei es, daß sonst gleiche Objekte in bezug auf eine Eigenschaft voneinander abweichen, sei es, daß sonst verschiedene Objekte in bezug auf eine Eigenschaft übereinstimmen.

Von konkreten Eigenschaftsvergleichungen sprechen wir dort, wo diese Vergleichsmerkmale sinnlich wahrnehmbar sind. So kann man geometrische Figuren vergleichen lassen in bezug auf ihre Größe, Eckenzahl, Lage im Raum, Farbe usw. Netschajeff (*89*) ließ zwei Bildertafeln vergleichen (MS. 183—184); die Objekte der einen Tafel fanden sich auf der anderen mit einigen Änderungen wieder: ein Hammer war hier nach rechts, dort nach links gerichtet; ein Elefant hat auf dem einen Bild einen herabhängenden, auf dem anderen einen aufgerichteten Rüssel, eine Schere ist hier geschlossen, dort geöffnet usw.

In diesen Tests ist die Vergleichstätigkeit auf sichtbare Eigenschaften gerichtet, wie auch im Alltagsleben wohl die optischen Vergleiche eine besonders große Rolle spielen. Um so mehr war die Idee Binets (*3*, *4*) zu begrüßen, eine Wahrnehmungseigenschaft zu prüfen, die in dieser Isolierung im Alltagsleben kaum je den Kindern begegnet, so daß hier die Prüflinge wirklich vor einer neuartigen Aufgabe standen. Es ist der Vergleich der Schwere von Gegenständen (MS. 182), die optisch völlig gleich sind (kleine schrotgefüllte Kästchen). Besonders hat sich dieser Test dort bewährt, wo über das paarweise Vergleichen hinaus zum Ordnen einer größeren Zahl von Objekten übergegangen wurde (s. S. 121).

Die konkreten Eigenschaftsvergleichungen können auch aus dem Gedächtnis vollzogen werden — was natürlich eine nicht unbeträchtliche Erschwerung bedeutet, da alles davon abhängt, ob das Kind sich die Gegenstände mit genügender Deutlichkeit vorstellen kann. So wird z. B. gefragt nach dem Unterschied von Fliege und Schmetterling, von Holz und Glas usw. Diese Fragen (MS. 186—189) gehören zum Grundstock des Binetschen Staffelsystems und sind für 8jährige bestimmt; neuerdings hat Kar-

städt (*73*) eine große Reihe solcher Unterscheidungsfragen auf ihre Brauchbarkeit hin geprüft.

Die abstrakten Eigenschaftsvergleichungen beziehen sich auf Eigenschaften, die nur dem Denken gegeben sind, z. B. auf ästhetische, moralische, logische, mathematische usw.[1]).

Schon bei den Definitionen hoben wir hervor, daß die Fähigkeit, sich und anderen von abstrakten Begriffen Rechenschaft zu geben, erst um das 12. bis 13. Lebensjahr zur Entwicklung kommt; dem entspricht es, daß auch um diese Zeit erst abstrakte Vergleichungen und Unterscheidungen möglich werden, z. B. zwischen Lüge und Irrtum, Geiz und Sparsamkeit usw. Solche Denkaufgaben eignen sich auch noch für höhere Altersstufen; so wurden in Hamburg schwierige Wortpaare, wie „Arbeit—Spiel", „Unterrichten—Erziehen" usw. mit Erfolg bei 14- und 15jährigen Mädchen, die ins Lehrerinnenseminar aufgenommen werden sollten, angewandt (MS. 189—190); es wurde dann verlangt, daß sie sowohl die Übereinstimmungen wie die Verschiedenheiten angeben, also eigentlich von jedem Begriff eine Definition mit Oberbegriff und differentia specifica liefern sollten.

Mathematische Eigenschaften lassen Moede und Piorkowski (*206*) vergleichen (MS. 185); sie bieten ihren (begabten 14jährigen) Prüflingen Vielecke mit verschiedener Eckenzahl und mehreren eingezeichneten Strichen; die gemeinsamen Merkmale (z. B., daß alle Figuren aus je sieben Strichen bestehen) sollen gefunden werden.

Die schwerste Vergleichsaufgabe liegt dort, wo nicht mehr Eigenschaften der einzelnen Gegenstände, sondern die zwischen mehreren Gegenständen bestehenden Beziehungen miteinander verglichen werden sollen, z. B. die Reihenfolge der Faulheit mehrerer Kinder mit der Reihenfolge der dazugehörigen Zeugnisse (Moede-Piorkowski). Die Hauptanwendung erfuhr diese Beziehungsvergleichung im sogenannten Analogietest (MS. 192 bis 194): Ein Begriffspaar mit bestimmter Beziehung wird gegeben,

---

[1]) Binets bekannter Test, zwei Gesichter auf ihre „Schönheit" zu vergleichen (MS. 186—187), gehört nicht hierher. Denn von den Gesichtern jedes Paares ist das eine ganz ausgesprochen häßlich, das andere ausgesprochen niedlich, so daß das Kind hier schon nach dem absoluten Eindruck, also ohne erst zu vergleichen, das „schöne" Gesicht zeigen kann. Der Test prüft also höchstens das Verständnis für den Ausdruck schön, nicht aber die Fähigkeit des Vergleichens.

von einem zweiten Begriffspaar nur das eine Glied; das andere muß auf Grund der entsprechenden Beziehung gefunden werden. Beispiele:

Gegeben wird: Apfel—Schale; Hase—; gefunden werden soll: Fell.
„ „ Hut—Kopf; Schuh—; „ „ „ Fuß.
„ „ Wasser—Schiff; Land—; „ „ „ Wagen.

Nach unseren Hamburger Erfahrungen ist der Test wegen der weitgehenden Abstufbarkeit der Schwierigkeit für sehr verschiedene Altersstufen brauchbar. Methodisch ist zu beachten, daß man nicht solche Wortpaare wählen darf, bei denen sich das zu ergänzende Wort auf rein assoziativem Wege leicht einstellen kann. Deshalb sind Analogien wie: „Vater—Sohn; Mutter—", und „Lanze—stechen; Gewehr—" ungeeignet, weil die Wörter „Tochter" bzw. „schießen" auch assoziiert werden können, ohne daß die analoge Beziehung im ersten Wortpaar aufgefaßt sein müßte.

20. Ordnungstests (MS. 194—209).

Als eine, wie es scheint, sehr aussichtsvolle Form des stummen Tests ist seit kurzem der Ordnungstest ausgebildet worden, der in drei Grundformen: der Stufenordnung, des Schemas und der Zeitordnung, auftreten kann. Die Elemente werden dem Prüfling ungeordnet dargeboten; er hat die Aufgabe, die richtige Ordnung herzustellen.

I. Stufenordnung.

Die Herstellung von Stufenordnungen geht aus der eben besprochenen Funktion der Vergleichung hervor; es muß nun nicht nur für zwei, sondern für mehrere Gegenstände sowohl ihre Übereinstimmung in den meisten Beziehungen wie auch ihre Verschiedenheit in bezug auf ein Merkmal erkannt, sodann dieses Merkmal als ein abstufbares aufgefaßt, endlich durch Vergleichung jedes Gliedes mit jedem anderen die richtige Abstufung gefunden werden.

Zum ersten Male wurde der Test durch Binet (*3*, *4*) angewandt (MS. 195)[1]). Fünf kleine, völlig gleich aussehende Kästchen, die aber durch verschiedene Füllung an Gewicht merklich verschieden sind, werden dem Prüfling vorgelegt. Der Prüfer sagt, daß diese Kästchen nach der Schwere zu ordnen seien, und macht es einmal vor; der Prüfling hat es dreimal nachzumachen. Natür-

[1]) Vgl. auch die Beschreibung des Tests und seiner Ergebnisse durch Bobertag (*51*).

lich darf der Gewichtsunterschied nicht so gering sein, daß der Ausfall von der peripheren Unterschiedsempfindlichkeit des Prüflings abhängen könnte; für die Intelligenzprüfung kommt es ja nur darauf an, daß der Prüfling die Aufgabe versteht, daß er das Merkmal der Schwere aus allen an den Kästchen wahrnehmbaren Merkmalen isolierend heraushebt, und daß er imstande ist, den Gedanken der aufsteigenden Schwere dauernd dem Vergleich der einzelnen Elemente zugrunde zu legen. Er muß ferner von selbst darauf kommen, daß es nicht genügt, je zwei Kästchen zu vergleichen und nebeneinander zu legen, sondern daß jedes neu einzuordnende Kästchen mit allen schon vorher geordneten verglichen werden muß. All dies sind ausgesprochene Denkanforderungen; und in der Tat bietet bei diesem Test nicht nur die schließliche richtige oder falsche Lösung, sondern auch das Verhalten des Prüflings während des Versuchs Anhaltspunkte zur Einschätzung seiner I.

Dieser Test ist für das Alter 9 Jahr angesetzt.

Mir schien nun das hier zum Ausdruck kommende Prinzip des „Ordnens nach einem leitenden Gesichtspunkt" so wertvoll, daß ich seine Erweiterung auf andere Ordnungsgesichtspunkte und zugleich seine Übertragung auf ältere Jahrgänge anstrebte.

Bezüglich der zu ordnenden Gegenstände kann man unterscheiden zwischen sensoriellen und begrifflichen Stufenordnungen. Die sensoriellen (MS. 195—198) können den verschiedensten Sinnesgebieten entlehnt werden; am leichtesten sind optische Reize herstellbar, nämlich Figuren verschiedener Größe, Helligkeit, Form und Farbe. Sehr leicht zu ordnen sind verschieden große Kreise, etwas schwerer verschieden helle, gleich große Quadrate, noch etwas schwerer Vielecke mit verschiedener Seitenzahl oder Sterne mit verschiedener Strahlenzahl.

Zur Herstellung einer begrifflichen Ordnung (MS. 199—204) wird eine Reihe von Begriffen, die in einem logischen Stufenverhältnis stehen, in ungeordneter Folge dargeboten, z. B.: Postbeamter, Mensch, Briefträger, Mann, Geldbriefträger, Beamter (nach dem Grad der Allgemeinheit zu ordnen). Oder: Monat, Stunde, Woche, Sekunde, Jahr, Tag, Minute (nach der Länge der Zeit zu ordnen) usw. Hier sind natürlich die verschiedensten Schwierigkeitsgrade herstellbar[1]).

[1]) Auch auf den S. 99 besprochenen Fernald-Jacobsohnschen Test (Ordnung von Verbrechen nach der Schwere) sei hier noch einmal hingewiesen.

Methodische Erschwerungen gegenüber dem Kästchentest sind auf verschiedene Weise möglich. Denkpsychologisch ist besonders wichtig, ob man den Gesichtspunkt, nach welchem geordnet werden soll, den Prüflingen darbietet (so bei Binet, der auffordert, die Kästchen „nach der Schwere" zu ordnen), oder ob man sie ihn selbst finden läßt. Es darf dann bei Vorlage der ungeordneten Elemente (Figuren, Begriffe) nur gesagt werden: Bringe diese in eine möglichst gute Ordnung. Nach erfolgter Lösung kann man Angaben über den gewählten Ordnungsgesichtspunkt verlangen.

Weiter kann die Ordnungsform variiert werden. Die eindimensionale lückenlose Reihe stellt den einfachsten Fall dar. Es folgt die lückenhafte Reihe: so wenn Vielecke mit aufsteigender Seitenzahl, aber unter Fortlassung des Siebenecks vorgelegt werden. Wird hier das fehlende Glied von selbst oder auf Befragen vermißt? — Sehr viel schwerer ist die Herstellung mehrdimensionaler Ordnungen.

II. Mehrdimensionale Ordnungen (Schemata) (MS. 200—204 und 247—249).

Hier haben wir wiederum sensorielle und begriffliche Ordnungen zu unterscheiden.

Die Aufgabe, Wahrnehmungselemente unter mehreren Vergleichsgesichtspunkten zu beachten und unter deren gemeinsamer Berücksichtigung zu ordnen, bereitet stets große Mühe. Wird z. B. eine Reihe von farbigen Papieren aus allen Gebieten der Farbenskala vorgelegt mit der Aufforderung, sie „möglichst gut" zu ordnen, so wird zwar eine Reihung nach der Ähnlichkeit relativ leicht gefunden; die Einsicht aber, daß die Wahl des Anfangs- und Endgliedes der Reihe willkürlich, die wahrhaft korrekte Lösung daher die Ordnung im Kreise ist, ist selbst bei Erwachsenen nur selten vorhanden.

Ein anderes hier anwendbares Ordnungsprinzip ist das Achsenkreuz. Diese Anordnung wird verlangt, wenn die Elemente zwei verschiedenen Stufenreihen angehören, die aber ein Glied gemeinsam haben; es werden z. B. verschieden große Quadrate von gleicher mittlerer Helligkeit und verschieden helle Quadrate von gleicher mittlerer Größe durcheinander gemischt vorgelegt. Hier muß erstens die Trennung beider Reihen erfolgen, dann die Zugehörigkeit des einen Gliedes zu beiden Reihen bemerkt werden, dann

das Mittel gefunden werden, um dieser doppelten Zugehörigkeit des Gliedes gerecht zu werden. Natürlich darf der Prüfling noch nicht in der Mathematik das Prinzip des Koordinatensystems kennengelernt haben. Ob das Selbstfinden dieses Prinzips wesentlich von mathematischer Spezialbegabung oder von allgemeiner I. abhängt, ist noch nicht untersucht worden. Zweifellos aber spielt die letztere hierbei eine nicht unbedeutende Rolle.

Bei gedanklichen Elementen kann die mehrdimensionale Ordnung die Form eines mehr oder minder verwickelten Begriffsschemas annehmen. Für solche Schematests (MS. 200—203; 247 bis 249) hat zuerst Giese Vorschläge gemacht (62a), die aber für Kinder und Jugendliche zu schwer sind, meist auch höhere Bildung voraussetzen. Sie gaben uns aber die Anregung, im Hamburger Laboratorium einfachere Schematests auszuarbeiten, die sich recht gut zu bewähren scheinen. Es wird den Prüflingen die textliche Aufzählung irgend einer wohlbekannten Gegenstandsgruppe (z. B. der Musikinstrumente, der Geldsorten, der Fahrzeuge) vorgelegt, die sie nun nach dem Prinzip der Über-, Neben- und Unterordnung schematisch darzustellen haben. An einem Beispiel wird ihnen klargemacht, was gemeint ist. Die Schwierigkeit ist hier besonders durch die Art des vorgelegten Textes abstufbar. Gibt man nur die Überschrift, z. B. „Fahrzeuge", so sind sowohl die zu ordnenden Elemente wie die Gesichtspunkte der Ordnung, wie endlich die schematische Darstellung selbst zu finden. Eine gewisse Erleichterung ist es schon, wenn die Elemente etwa in alphabetischer Reihenfolge geboten werden. Wir gingen in der Erleichterung noch viel weiter, und da zeigte sich, daß auch so noch 14- und 15jährige Prüflinge vor manchen intellektuellen Schwierigkeiten standen. Wir boten im Text nicht nur die Elemente, sondern sogar die Gesichtspunkte der Einteilung dar, z. B.: „Fahrzeuge gebraucht man auf dem festen Lande, auf dem Wasser und in der Luft. Auf dem festen Lande werden die Fahrzeuge teils durch Menschenkraft (Karren, Fahrräder), teils durch Tierkraft (Lastwagen, Droschken) bewegt . . ." Die Prüflinge haben also lediglich eine Übertragung des Sprachlichen ins Logisch-Schematische vorzunehmen. Aber auch die hierzu nötige Umdenkung und die Loslösung von den Zufälligkeiten der textlichen Formulierung gelingt nicht ohne weiteres.

Charakteristisch ist z. B. folgender Fehler: Im Text steht bei den Landfahrzeugen (siehe oben) erst die Kraft und dann der Name der Fahrzeuge. Bei

den Wasserfahrzeugen ist die Reihenfolge im Text umgekehrt („Segelboote, die der Wind vorwärts treibt") — nun halten sich manche Kinder so sklavisch an den Text, daß sie diese Ungleichmäßigkeit auch in ihr Schema übertragen — ein unverkennbares Zeichen schwächerer I.

Als Umkehrung dieses Schematests habe ich ferner das Lesen eines Schemas vorgeschlagen (MS. 204): es soll z. B. ein einfaches genealogisches Schema in sprachliche Formulierung umgesetzt werden. Beispiele und Erläuterungen müssen stets so eingerichtet sein, daß keinerlei Vorkenntnisse über den Gebrauch schematischer Darstellungen vorausgesetzt zu werden brauchen. Nach einigen Vorproben, die wir angestellt haben, ist dieser Test leichter als der oben beschriebene, daher bei jüngeren Jahrgängen (etwa 12—13jährigen) anwendbar.

III. Zeitordnungen.

Die einzelnen Phasen eines in sich zusammenhängenden Ereignisses werden in Sätzen oder in Stichworten oder in Bildern dargestellt und durcheinander gemischt. Bei der Aufforderung zur Ordnung braucht nicht mehr, wie bei den logischen Ordnungen, ein Ordnungsgesichtspunkt dargeboten oder gesucht zu werden; denn es erscheint selbstverständlich, daß das Nacheinander der geordneten Elemente dem Nacheinander der Ereignisphasen entsprechen muß. Die hergestellte Ordnung soll dann eben die Geschichte des Ereignisses erzählen.

Mit Stichworten arbeiten die im Hamburger Laboratorium ausgearbeiteten Tests (MS. 205—207). Beispiel einer solchen ungeordneten Reihe: Arzt, Fußballspiel, Heilung, Verband, Beinbruch, Besserung, Sturz. Es zeigte sich, daß 10jährigen Kindern die straffe Durchführung der kausal bestimmten Abfolge durchaus nicht immer gelingt; zuweilen kreuzen andersartige Motive die richtige Ordnung. Für ältere Prüflinge wird der Test dadurch geeignet, daß man die Stichworte zweier unabhängiger Geschehnisfolgen durcheinander mischt und die Herstellung beider Reihen fordert.

Der Leipziger Lehrerverein (*197*) benutzt statt der Stichworte die vollen Sätze einer Geschichte, die auf einzelne Zettel geschrieben sind und in die richtige Reihenfolge gebracht werden müssen (MS. 239).

Als sehr ergiebig erwies sich die Ordnung einer Bilderfolge (MS. 207—209). Die Verwendung von Bilderbogen be-

gegnete uns schon einmal (S. 93); dort wurden die vom Text befreiten Bilderbogen in der richtigen Abfolge vorgelegt, um vom Prüfling beschrieben zu werden. Hier werden die Bilderbogen zerschnitten vorgelegt, und der Prüfling hat selbst die richtige Anordnung der durcheinander gemischten Einzelbilder zu finden.

Die Bilderbogen müssen auf ihre Eignung zu diesem Versuch besonders geprüft werden; finden sich doch oft in ihnen Darstellungen von Phasen, die nicht eindeutig an eine bestimmte Stelle der Reihe gebunden sind; solche Teilbilder sind fortzulassen. Jedes der vorgelegten Bilder muß sich zweifellos in eine zeitliche Handlungsabfolge einfügen. Andererseits kann man durch Fortlassung weiterer Phasen die Aufgabe erschweren, indem nun wichtige Abschnitte der Handlung kombinatorisch ergänzt werden können. Hierdurch sowohl wie durch die Auswahl verschieden schwerer Bilderbogen ist die Schwierigkeit mannigfach abstufbar.

Ordnungsversuche mit Bilderbogen sind bereits vor Jahren im Breslauer psychologischen Seminar angestellt, dann im Hamburger psychologischen Laboratorium fortgesetzt worden. Eine Reihe von Bilderbogen wurde ausgeprobt; sie sind für sehr verschiedene Altersstufen (zwischen 9 und 14 Jahren) brauchbar. Endgültige Ergebnisse liegen noch nicht vor. Unabhängig davon hat der belgische Forscher Decroly (57) den gleichen Test ausgebildet; er verwandte zum Teil so leichte Bilderfolgen, daß sie schon von 7- und 8jährigen gelöst wurden. Normale und schwachsinnige Kinder zeigten sehr starke Leistungsunterschiede.

Bei allen Ordnungstests ist das Verhalten während der Arbeit psychologisch weit wichtiger als das schließliche Endergebnis. Wenn man Kinder von verschiedener Befähigung beim Ordnen der fünf Gewichtskästchen oder beim Legen einer Bilderfolge beobachtet, so gewinnt man eine ganze Reihe von höchst fesselnden Einblicken in die Art ihrer Denktätigkeit. Nähere qualitative Studien, die sehr zu empfehlen sind, werden hierbei nicht nur Grade, sondern auch Typen der I.-Leistung hervortreten lassen: ob mehr systematisch oder sprunghaft gearbeitet wird, ob erst viele Möglichkeiten durchgeprüft werden oder sofort mit intuitiver Sicherheit der rechte Ordnungsgesichtspunkt gewonnen wird, ob Selbstkorrekturen in größerem Umfang vorkommen, ob die Aufmerksamkeit immer nur auf wenige Glieder der Reihe konzentriert wird oder möglichst viele Elemente zu überschauen sucht.

Aus dem eben genannten Grunde werden Ordnungstests vor allem bei Einzelprüfungen in Betracht kommen, bei denen über den Verlauf der Tätigkeit von einem Beobachter Aufzeichnungen und nachher vom Prüfling über seine eigenen Erlebnisse während der Arbeit Aussagen gemacht werden können. Immerhin ist es auch möglich, manche Ordnungstests, so vor allem das Ordnen von Begriffen und von Stichworten, für die Anwendung bei Massenprüfungen umzuformen.

21. Zuordnungstests (MS. 209—212).

„Zuordnung" ist diejenige Tätigkeit, die zu je einem dargebotenen Begriff aus einer Reihe anderer den sinnvoll dazugehörigen Begriff findet; in dieser Tätigkeit ist Kombinieren und kritisches Wählen verbunden. Hierfür hat A. Franken (*61*) einen stummen Test ausgearbeitet. Das Reizmaterial besteht in zwei Tafeln mit je 52 Bildchen. Mit jedem Bildchen der einen Tafel steht eines der zweiten im sinnvollen Zusammenhang, so Spinne und Spinnennetz, Stiefel und Stiefelknecht, Gewicht und Wage, Schmetterling und Raupe usw. Nach F.s Befunden steigt die Fähigkeit, richtige Kombinationen zu finden und falsche zu vermeiden, rasch in den Altersstufen von 10 bis zu 12 Jahren; bedeutende Unterschiede zeigten sich auch bei verschiedenen Begabungsstufen gleichaltriger Kinder.

Auch die Zuordnung von Bild zu Text wäre verwendbar. Es gibt Münchener Bilderbogen, welche Illustrationen zu Sprichwörtern enthalten. Werden von diesen Bildern die Unterschriften entfernt, andererseits die Sprichwörter in einer alphabetisch geordneten Liste vorgelegt, so kann nunmehr die Zuordnung verlangt werden.

VII. Kapitel.

Prüfsysteme zur Messung des Intelligenzgrades.

Wie schon im Kapitel V erwähnt, kann man sich bei der praktischen Anwendung der IP. fast nie mit einem einzelnen Test begnügen; es müssen vielmehr Sérien von Prüfmitteln zusammengestellt werden, welche die I. von recht verschiedenen Seiten her abstecken. Hierbei muß eine weise Auswahl getroffen werden, indem aus der unübersehbaren Fülle möglicher Tests solche mit

hohem Symptomwert, allgemeiner Anwendbarkeit und objektiver Messungsfähigkeit herausgegriffen und so zusammengestellt werden, daß die Hauptseiten der I. durch charakteristische Äußerungen vertreten sind. Es müssen ferner Wertungsmethoden angewandt werden, welche die Herstellung eines Resultantenwertes als einer Gesamtformel für den I.-Grad des Prüflings erlauben.

In der Zusammenstellung und Erprobung von Prüfserien waren die Psychiater vorangegangen. Auch haben manche Pädagogen, insbesondere Praktiker der Heilerziehung, hier und da aus eigener Initiative Prüfungen des Geistesstandes ihrer Zöglinge erdacht. Aber in dem Maße, als man sich der eben angedeuteten Forderungen bewußt wurde, die an solche Prüfserien zu stellen sind, erkannte man auch die große Schwierigkeit dieser wissenschaftlichen Aufgabe; und es erschien nicht mehr möglich, daß sie gleichsam nebenbei von Psychiatern und Erziehern gelöst werden könnte. Die Psychologie mußte nun die Hauptarbeit übernehmen, und sie ist seit etwa zwei Jahrzehnten daran tätig, ohne daß ein Abschluß auch nur annähernd zu erblicken ist. Immerhin ist auch das bisher Erreichte schon bemerkenswert genug, insbesondere wenn man die Jugend unserer Wissenschaft und die geringe Zahl ihrer Mitarbeiter in Betracht zieht.

In dieser Arbeit gibt es neben erfreulichen Fortschritten und neuen Anwendungsmöglichkeiten auch manches Unerwünschte: Voreiligkeiten, Umwege und Rückschläge. Aber solche Kinderkrankheiten werden überwunden werden; und zu ihrer schnelleren Überwindung können vielleicht gerade solche Zusammenfassungen und Selbstbesinnungen beitragen, wie dieses Buch sie zu geben wünscht.

Die bisher aufgestellten Prüfungsserien lassen sich in drei Gruppen teilen. Gruppe 1 umfaßt die zu psychiatrischen Zwecken erdachten Systeme, Gruppe 2 das Staffelsystem von Binet-Simon, das ein absolutes I.-Maß erarbeiten will, Gruppe 3 solche Testreihen, die für ganz bestimmte Probleme: Untersuchungen der Kinder einer Schulklasse, Auslese begabter Kinder usw. zusammengestellt sind und sich mit einer relativen Messung der I. begnügen.

Das vorliegende Kapitel hat es nur mit festen Prüfsystemen zu tun, die nicht von Fall zu Fall veränderlich sind, sondern in ihrem Gesamtaufbau eindeutig einem bestimmten I.-Bild oder -Maß

zugeordnet sind. Es handelt sich hierbei in erster Linie um die Staffelmethodik der Gruppe 2; vorausgeschickt wird eine kurze Erörterung über Gruppe 1. Dagegen können die beweglichen, für Sonderzwecke bestimmten Prüfserien erst dort dargestellt werden, wo eben die besonderen Probleme, für die sie geschaffen worden, zur Behandlung kommen, also im dritten und vierten Hauptteil (vgl. insbesondere Kap. X und Kap. XIII).

1. Psychiatrische Prüfreihen.

Von psychiatrischer und nervenärztlicher Seite sind seit Rieger (*94*) zahlreiche Testreihen angewandt worden; doch da diese in erster Reihe dazu bestimmt waren, den Intelligenzgrad erwachsener Patienten zu prüfen, genügt an dieser Stelle eine kurze Erwähnung. Gewöhnlich ist derartigen Serien ein psychologisches Schema zugrunde gelegt, das freilich von Forscher zu Forscher sehr wechselt. Es seien beispielshalber drei solcher Listen zitiert.

Sommer (*110*) bespricht in einem Vortrag über die Methoden der IP. der Reihe nach die Prüfmittel für folgende Teilprobleme: Beziehung (des Gedächtnisses, der Schulkenntnisse des Rechenvermögens, der Assoziation zum Verstand, ferner die Aufmerksamkeit, die Auffassungsfähigkeit, Vollständigkeit der Komplexe, Analyse der Komplexe, Ergänzung von Komplexen, mechanischen Verstand (Findigkeit), konstruktiven Verstand, logische Unter- und Überordnung, Kausalitätsbegriff, intellektuelles Interesse, Verständnis für die Umwelt.

Ziehen (*126*) teilt in seinen „Prinzipien und Methoden der IP." nach folgenden Gesichtspunkten ein: Retention, Vorstellungsentwicklung und Vorstellungsdifferenzierung (Generalisation, Isolation und Komplexion von Vorstellungen), Reproduktion und Kombination, und beschreibt bei jeder Abteilung die zahlreichen in seiner Klinik gebräuchlichen Frage- und Prüfungsmethoden

Cimbal (*54*) zählt in seinem nervenärztlichen Taschenbuch folgende Gebiete von geistigen Fähigkeiten auf, für deren jedes er eine Reihe von Prüfmitteln vorschlägt: Arbeitsvermögen, Auffassungsvermögen, optische Prüfungen, Denkvermögen (zu prüfen durch Bilder und durch sprachliche Aufgaben), Urteilsvermögen.

Wenn man diesen und anderen Serien der Psychiater die Vielseitigkeit nicht absprechen kann, so unterliegen sie doch in anderer Hinsicht manchen Bedenken. Alle Serien machen nämlich den Eindruck, als ob die Auswahl der Tests mehr durch Zufall und Willkür als auf Grund einer wirklichen Eichung erfolgt wäre. Meist war ein apriorischer Gedankengang bestimmend: daß eine gewisse Fähigkeit (z. B. die Retention oder die Kombination), welche als

zur I. gehörig angenommen wurde, durch ein gewisses Prüfungsmittel getroffen werde. Ob gerade dieser Test vor so und so vielen anderen durch Schärfe, Konstanz und Bedeutsamkeit der durch ihn erzielten individuellen Werte ausgezeichnet sei, diese Frage ist nur ganz selten zum Gegenstand einer eigenen Voruntersuchung gemacht worden. Die Folge dieser Zufallsauswahl ist dann aber auch die mangelnde Übereinstimmung in den Testserien verschiedener Forscher; jede psychiatrische Klinik besitzt ihre Sondermethodik der IP., jeder Nervenarzt, jeder Hilfsschularzt wählt sich seine Tests nach privater Liebhaberei aus; und so ist eine wirkliche Vergleichung, Nachprüfung und Vereinigung der Ergebnisse verschiedener Untersuchungen bisher nur in geringem Maße möglich gewesen.

Endlich leiden die üblichen psychiatrischen Testreihen daran, daß ihnen ein Prinzip der Zusammenfassung der Ergebnisse fehlt. Daß es möglich ist, die I. eines Individuums in ihrer Gesamtheit zu bewerten, erkennen die Psychiater an, indem sie die Prädikate „beschränkt zurechnungsfähig", „debil", „imbezill", „idiotisch" verwenden; aber sehen wir zu, wie sie im individuellen Fall von den Befunden ihrer Testreihen zu dem Gesamturteil kommen, so klafft eine Lücke. Das Mosaik der Testergebnisse ist und bleibt nur Rohstoff; kein methodischer Grundsatz, sondern subjektive Würdigung jener Befunde, Intuition und Routine geben die Entscheidung für das Gesamtzeugnis, das der I. beigelegt wird. Darin liegt in gewissem Sinne ein Vorzug; denn die — fast künstlerische — Gabe der intuitiven Einfühlung und des nacherlebenden Verstehens ist gerade für den Psychiater unentbehrlich. Wird aber alles dieser Fähigkeit allein überlassen, so ist auch ein starker Nachteil vorhanden; denn jedes Urteil bleibt dann subjektiv, ist nicht zu kontrollieren und zu verallgemeinern. Daher ist das Bestreben berechtigt, daß zum mindestens neben jene intuitive Diagnose auch ein Verfahren der objektiven Gradbestimmung trete.

Als ein von psychiatrischer Seite stammender Versuch einer solchen Systematik ist die Methode der psychologischen Profile zu erwähnen, die von dem Russen Rossolimo (*99*, *100*) ausgebildet wurde. Sie soll das Individualbild jedes Prüflings in die Form eines optischen Diagramms (des „Profils") bringen. Für 10 verschiedene seelische Funktionen werden Tests angewandt; innerhalb jeder Testgruppe ist eine Abstufung der Leistungen mit

den Wertziffern 1—10 vorgesehen. Nun erhält der Prüfling für jede Testgruppe nach dem Leistungswert eine höhere oder geringere Ordinate, und die Verbindung der Ordinatenenden liefert dann das „Profil". Aus einem solchen kann man dann ablesen: X hat im Wortgedächtnis eine hohe, in der Assoziationsfähigkeit eine mittlere, in der Kombination eine geringe Leistungsfähigkeit usw. Typisch verschiedene Profilformen stellen typisch verschiedene Verteilungen der Leistungsfähigkeit im Individuum und damit typisch verschiedene Strukturbilder dar. Diese Abweichungen sind insbesondere für die einzelnen Arten psychischer Defekte kennzeichnend.

Die Methode ist in ihrer ursprünglichen Form wohl vornehmlich zur Prüfung Erwachsener bestimmt; auch beschränkt sie sich nicht auf intellektuelle Funktionen, sondern zieht auch andere mit hinein, so daß sie nicht eigentlich in den Rahmen dieses Buches gehört. Etwas willkürlich mutet zuweilen die Zuordnung bestimmter Tests zu bestimmten Seelengebieten an.

Für unsere Zwecke sind zwei Abwandlungen der Methode wichtiger, die neuerdings hervorgetreten sind.

Rossolimo (*101*) selbst hat eine Vereinfachung vorgeschlagen, die sogenannte „kurze Methode", die zur ersten vorläufigen IP. von Kindern dienen soll. Die Reihe besteht aus 27 Versuchen, in denen geprüft wird: die Aufmerksamkeit (Durchstechen von Punkten auf einem Karton), die Merkfähigkeit für einfache geometrische Gebilde, das Gedächtnis für Bilder und Sätze, die Auffassungsfähigkeit für Bilder (Einzelbilder, Bilderserien, absurde Bilder); die kombinatorische Fähigkeit (Zusammensetzen von Bildern und Figuren); der mechanische Sinn (Findigkeit für eine einfache mechanische Aufgabe), die Einbildungskraft (Ergänzung eines unfertigen Bildes); Beobachtungsfähigkeit (für Verschiedenheit geometrischer Figuren). Auch hier erscheint die Auswahl und Zuordnung der Tests zu den bestimmt benannten Seelenfunktionen einigermaßen willkürlich.

Verzichtet man aber auf die Gliederung nach Funktionen und begnügt man sich mit dem Gesamtergebnis aus den 27 Tests, so scheint das Verfahren für den begrenzten oben genannten Zweck brauchbar zu sein. Es lieferte nach den Ergebnissen von Rabinovitsch (*91*) sehr deutliche Unterschiede zwischen fortgeschrittenen und rückständigen Schulkindern, die aus der Normalschule stammten, noch größere zwischen normalen, debilen und imbezillen

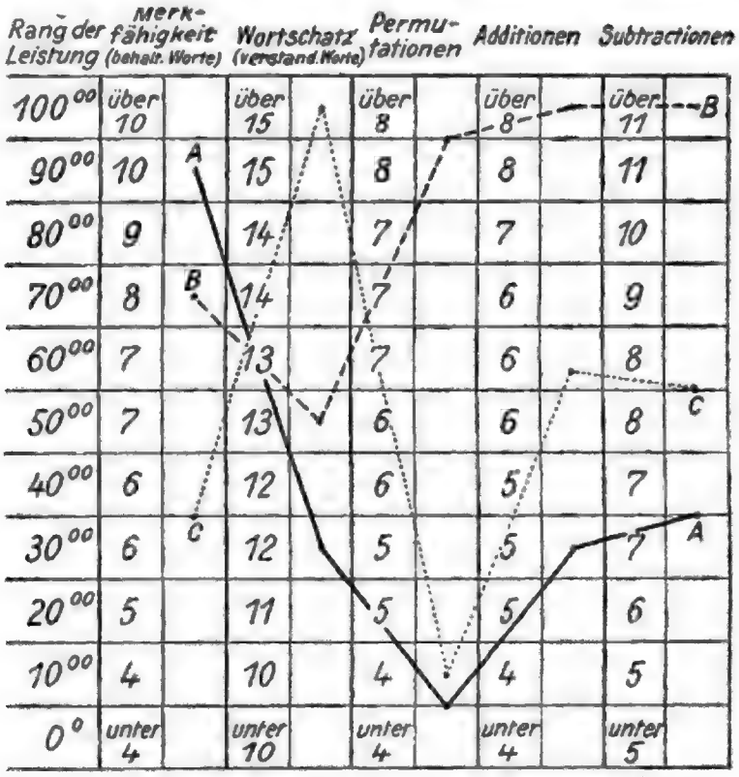
Kindern; die Ziffern der durchschnittlich gelösten Tests betrugen bei diesen drei Gruppen: 26,7, 19, 11. Die Altersunterschiede traten den Begabungsunterschieden gegenüber zurück. Die zur Prüfung nötige Zeit (10—30 Minuten) wuchs mit dem Grade der Zurückgebliebenheit.

In einem gewissen Gegensatz zu dieser Vereinfachung steht die Verfeinerung der Methode, die Claparède (*56*) vorschlägt und die eine genauere Nachprüfung auf ihre Brauchbarkeit verdient. Cl. bemängelt die Willkür, mit der bei Rossolimó jede Gruppe von Testleistungen nach den Rangstufen 1—10 abgestuft ist. Diese Bewertung geht nach beliebigen Maßstäben, z. B. nach der Anzahl der behaltenen Worte oder der gelieferten Assoziationen vor, und es ist keine Gewähr dafür vorhanden, daß die Rangstufe 5 bei dem Gedächtnis die gleiche mittlere Leistungsfähigkeit bedeute, wie dieselbe Rangstufe bei der Aufmerksamkeit oder der Kombination. Dies aber ist die Grundvoraussetzung der Profilidee; wagerechter Lauf des Profils muß Gleichwertigkeit der Leistungen, abfallender Lauf muß Minderwertigkeit des zweiten Leistungsgebiets gegenüber dem ersten bedeuten.

Hier setzt nun Claparèdes Besserungsvorschlag ein. Vergleichbar sind die Leistungsgrade eines Individuums auf verschiedenen Gebieten offenbar nur dann, wenn sie sämtlich auf die Leistungen der Gruppe bezogen werden, denen das geprüfte Individuum angehört. Denken wir uns der Einfachheit halber diese Gruppe aus 100 Personen bestehend, die alle mit denselben Tests geprüft und für jeden Test in eine Rangordnung gebracht sind, so kann man für das Individuum X angeben, welche Rangnummer es in jedem Einzeltest erhalten hat: die Verbindung dieser Rangnummern ergibt dann das Profil. Um das Schema zu vereinfachen, schlägt Claparède vor, für jeden Test die Rangordnung als 100prozentige darzustellen und sie in 10, gleich viele Individuen enthaltende Stufen zu gliedern. Er verwandte z. B. einen Merkfähigkeitstest mit 15 einmal vorgesprochenen Worten und fand, daß die 10% schlechtesten Versuchspersonen weniger als vier Worte behielten, die nächsten 10% der Prüflinge vier Worte, die folgenden 10% fünf Worte . . ., die zweitbesten 10% zehn Worte, die besten 10% mehr als zehn Worte. Damit ist die Leistung in 10 Rangstufen geteilt und für jedes mit dieser Gruppe vergleichbare Individuum läßt sich nun feststellen, ob seine Merkfähigkeit für die Worte eine 10prozentige, 20prozentige . . ., 50-

prozentige . . ., 100prozentige ist. Geschieht das gleiche für jeden anderen Test, so kann X mit seiner Individualleistung überall in analoger Weise eingeordnet werden, und das Gesamtprofil seiner Leistungen erhält einen eindeutigen Sinn.

Als Probe der Claparèdeschen Methode geben wir ein von ihm gebrachtes Diagramm wieder. Es bezieht sich auf fünf Tests: Merkfähigkeit für 15 Worte, Wortschatzprüfung (wieviel von 25 Worten der französischen Sprache wurden sinnvoll aufgefaßt?), Permutation der Buchstaben a b c d, Additions- und Subtraktionsaufgaben. Die Ziffern in jeder Testkolumne geben an, welche Anzahl richtiger Elemente (Worte, Permutationsformen, gerechnete Aufgaben) bei 12-jährigen Knaben in jedem Test als 10prozentige, 20prozentige . . . , 90prozentige Leistung angenommen sei. Für drei individuelle Prüflinge sind dann die Profile eingezeichnet, deren Sinn nun ohne weiteres verständlich ist. A hat seine Stärke im Wortgedächtnis, B im Rechnen, C im Wortverständnis.



Die Auswahl der zur Vergleichung gestellten Tests ist in dem Beispiel noch dürftig und willkürlich; auch muß man künftig Tests wählen, bei denen sich die Leistungen besser differenzieren als hier, wo allzu oft gleiche Testergebnisse auf verschiedene Rangstufen kommen; doch wird durch die Mängel dieser ersten nur probeweisen Anwendung der Wert der Methode an sich nicht betroffen.

2. Die Methode der Altersstaffelung nach Binet-Simon (BS.): Das Prüfungsverfahren.

Bahnbrechend für die Mitarbeit der Psychologie an der Aufgabe der IP. wurde der Versuch von Binet (*3*), ein nach Alters-

stufen gestaffeltes Prüfungssystem aufzustellen. Unterstützt wurde er hierbei durch den Arzt Simon. Den Anstoß gab ihm der Wunsch, die Überweisung schwach befähigter Kinder an die Hilfsschulen von einer exakten Feststellung ihres Rückständigkeitsgrades abhängig zu machen; aber im weiteren Fortschritt der Arbeit zeigte es sich, daß die Methode neben ihrer diagnostischen Bedeutung für Abnorme auch wertvoll als Forschungsmittel auf dem Gebiet der normalen I. ist, und daß sie insbesondere durch ihre leichte massenstatistische Anwendung wichtige Ergebnisse über die Verteilungs- und Abhängigkeitsbeziehungen der I. zu liefern vermag. Ich halte es sogar nicht für unmöglich, daß diese differentiellpsychologische Bedeutung in Zukunft noch immer stärker gegenüber der diagnostischen in den Vordergrund treten wird.

Die Forderungen, mit denen Binet und Simon an das zu schaffende Verfahren (weiterhin BS.-Methode genannt) herantraten, waren etwa die folgenden:

Es sollte für jede Altersstufe der Kindheit eine Testserie gefunden werden, deren Lösung eben gerade für Kinder dieses Alters als normal und charakteristisch gelten kann; die Tests sollten relativ unabhängig von äußerlichen und zufälligen Bedingungen, insbesondere von den Schulkenntnissen, sein, damit das Ergebnis möglichst rein die wirkliche geistige Veranlagung des Kindes zum Ausdruck brächte; sie sollten eine möglichst gleichmäßige Anwendung unter den verschiedensten Bedingungen der Nation, der Sprache, der Bildung erlauben; sie sollten leicht ausführbar, nicht an Laboratorium und Präzisionsapparate gebunden sein, den Prüfling nicht zu lange Zeit in Anspruch nehmen, nicht belästigen und ermüden, zugleich aber eine solche Exaktheit besitzen, daß die von verschiedenen Forschern vorgenommenen Untersuchungen miteinander verglichen und aneinander gemessen werden könnten; sie sollten endlich die Herstellung eines Resultantenwertes für jeden Prüfling erlauben, der als Maßzahl seiner Gesamt-I. gelten kann.

Es sieht zunächst so aus, als ob die Erfüllung so verschiedenartiger Ansprüche auf unüberwindliche Schwierigkeiten stoßen müßte. Fehlte doch vor allem jede Vorkenntnis darüber, welche intellektuellen Leistungen man auch nur ungefähr bei einem Kinde bestimmten Alters erwarten dürfe. Man frage einmal einen Lehrer oder eine Persönlichkeit, die sich jahrelang mit Kindern verschiedenen Alters beschäftigt hat: von welcher Altersstufe man wohl verlangen könne, daß der Unterschied zwischen zwei genannten

Objekten (z. B. Holz und Glas) richtig angegeben werde, und welche Altersstufe zur Angabe des Unterschieds zwischen abstrakten Begriffen (z. B. Lüge und Irrtum) befähigt sei — der Gefragte würde schweigen oder ziemlich blind darauflos raten müssen. So war also hier völliges Neuland zu bearbeiten. Dazu noch die anderen obengenannten Bedingungen, die zum Teil nur schwer miteinander vereinbar schienen: Unabhängigkeit von Schulkenntnissen, allgemeine leichte Anwendbarkeit, Kürze der Prüfung, Exaktheit, Berechnungsmöglichkeit — kein Zweifel, daß hier eines der schwersten Probleme vorlag, welches die angewandte Psychologie sich bisher überhaupt gesetzt hat.

Das Temperament und der Geist Binets ermöglichten es ihm, den kühnen Wurf zu wagen und sich dort, wo die wissenschaftlichen Vorarbeiten zu einer befriedigenden Überwindung der Schwierigkeiten noch lange nicht ausreichten, sich auf seine Intuition zu verlassen. Vielleicht wäre ohne dieses kecke Zugreifen überhaupt niemals der Anfang gemacht worden; und deshalb dürfen wir trotz aller Unvollkommenheiten der ersten Ausführung das hohe Verdienst dieser Tat würdigen. Freilich: es blieb — und bleibt — für die Nachfolger noch viel an dem System zu modeln und zu bessern; und schlimm ist es, daß geistlose Nachbeter aller Orten (namentlich in Amerika) wegen der scheinbar leichten Anwendbarkeit der Methode aus dem, was ein erster Versuch war, sofort ein starres und fertiges Allerweltswerkzeug zur Messung kindlicher I. zu machen suchten. Aber der einfache Grundgedanke der Altersstaffelung ist doch so wertvoll, daß er sich gegenüber Mißbräuchen und Irrungen durchsetzen wird. Denn wenn wir schon längst gewohnt waren, bei Erwachsenen die geistige Rückständigkeit durch Vergleich mit einer angenommenen Normal-I. festzulegen, so brauchen wir auch Entsprechendes bei Kindern; hier ist aber das Problem der „Normal-I." ein ganz andersartiges; denn deren Leistungsgrad steigt mit dem Alter und wir haben daher für jede Altersstufe einen anderen Kanon nötig, an dem wir die etwaige geistige Minder- (oder Höher-)Wertigkeit eines Kindes messen können. Dieser Forderung suchten Binet und Simon durch ihre Teststaffeln zu entsprechen.

Nach mehrfachen empirischen Voruntersuchungen darüber, welche Tests für bestimmte Lebensalter als normal zu betrachten seien, haben B. und S. im Jahre 1908 die erste Gesamtaufstellung ihres Testsystems veröffentlicht (*3*), welches Serien für die Alters-

stufen von 3 bis zu 13 Jahren enthielt. Für jede Altersstufe sind 5—7 Tests vorgesehen. Eine durchgesehene Fassung erschien 1911 (*4, 49*); in ihr sind manche Tests verändert, manche einer anderen Altersstufe zugewiesen und die Testanzahl für jede Altersstufe gleichmäßig auf 5 gebracht. Das System von 1911 enthält statt der Tests für 11-, 12- und 13jährige solche für 13jährige, 15jährige und Erwachsene.

Im gleichen Jahre starb Binet und nun ging die weitere Arbeit am Ausbau des Staffelsystems auf andere Länder über. Hier galt es, einerseits offenbare Schwächen und Lücken des bisherigen Systems zu beseitigen, andererseits diejenigen, übrigens ziemlich geringfügigen, Änderungen vorzunehmen, die durch die verschiedenen nationalen Verhältnisse bedingt waren. In Deutschland wurde diese Bearbeitung und Erprobung des Systems durch Dr. Bobertag (*51*) in Breslau vorgenommen, der 1912 eine ausführliche Darstellung veröffentlichte. Diese Fassung des Systems ist für Deutschland ziemlich allgemein in Gebrauch genommen worden, obgleich auch sie nicht von Mängeln frei ist. Eine sehr große Verbreitung fand die Methode in Amerika, wo die „Binetisten" bei ihrer Anwendung nicht selten die nötige Kritik vermissen ließen[1]). Von den mancherlei Bearbeitungen seien hier zwei erwähnt: die von Goddard (*63*), die sich sehr eng an Binet anschließt, und die „Stanford-Revision" von Terman (*118*), welche auf Grund sorgfältiger Massenprüfungen eine Neueichung der Tests vornimmt und hierbei zum Teil zu anderen Alterszuweisungen kommt als Binet. Terman geht auch über die obere Grenze der Altersstufen hinaus, indem er Teststaffeln für die Jahrgänge 14, 16, 18 hinzufügt. In ähnlicher Weise wie Terman hat Jaederholm (*71*) in Schweden exakte Kontrolluntersuchungen angestellt und eine zutreffendere Anordnung der Tests vorgeschlagen. Die Bearbeitungen der beiden letztgenannten bedeuten zweifellos einen großen Fortschritt gegenüber den früheren Testsystemen, und es wäre sehr zu wünschen, wenn wir auch die für Deutschland nötige Neubearbeitung bald bekämen; sie setzt freilich Massenprüfungen von großem Umfang und strenger Exaktheit voraus. Der Vorschlag einer neuen Skala, den Meumann (*9*) machte, kann aus später zu besprechenden Gründen nicht als das zu erhoffende deutsche Staffelsystem gelten;

---

[1]) Es scheint, daß dort das „Binetisieren" geradezu zu einem Unfug ausgeartet ist. Vgl. die scharfe Kritik von Haberman (*65b*).

er ist auch meines Wissens nicht praktisch zur Anwendung gekommen.

Da die fünf wichtigsten Fassungen des Staffelsystems in der Methodensammlung (*8*, III) synoptisch zusammengestellt sind, so kann hier ein Abdruck der jetzt in Deutschland verbreitetsten Fassung von Bobertag genügen[1]).

Dreijährige.

1. Mund, Augen, Nase zeigen.
2. Nachsprechen 6silbiger Sätze.
3. Nachsprechen von 2 Ziffern.
4. Bildbetrachtung: Aufzählung.
5. Angabe des Familiennamens.

Vierjährige.

1. Angabe des Geschlechts.
2. Benennen vorgezeigter Gegenstände.
3. Nachsprechen von 3 Ziffern.
4. Vergleich zweier Linien.
5. Vergleich zweier Gewichte.

Fünfjährige.

1. Nachsprechen 10silbiger Sätze.
2. Abzeichnen eines Quadrates.
3. Definition konkreter Gegenstände durch Zweckangabe.
4. Vier Pfennige abzählen.
5. Nachsprechen von 4 Ziffern.

Sechsjährige.

1. Bildbetrachtung: Beschreibung.
2. Nachsprechen 16silbiger Sätze.
3. Ästhetischer Vergleich.
4. Geduldspiel.
5. Ausführen dreier auf einmal gegebener Aufträge.

Siebenjährige.

1. Lücken in Bildern erkennen.
2. Rechts und links unterscheiden.
3. Abzeichnen eines Rhombus.
4. Nachsprechen von 5 Ziffern.
5. Kenntnis der Münzen von 1 Pf. bis 1 M.

Achtjährige.

1. Vergleich zweier Gegenstände aus dem Gedächtnis.
2. Rückwärtszählen von 20—1.
3. Benennung der 4 Hauptfarben.
4. Angabe eines Hauptpunktes aus einer eben gelesenen Zeitungsnotiz.
5. Drei leichte Verstandesfragen.

Neunjährige.

1. Definition durch Oberbegriffe.
2. Ordnen von Gewichten.
3. 80 Pf. auf 1 M. herausgeben.
4. Tagesdatum.
5. Bildbetrachtung: Erklärung mit Hilfe unterstützender Fragen.

Zehnjährige.

1. Angabe von 6 Erinnerungen aus einer gelesenen Zeitungsnotiz.
2. Kenntnis aller Münzen.
3. Mit drei Worten zwei Sätze bilden.
4. Nachsprechen 26silbiger Sätze.
5. Nachsprechen von 6 Ziffern.

Elf- und Zwölfjährige.

1. Kritik absurder Sätze.
2. Mit 3 Worten einen Satz bilden.
3. Drei schwere Verstandesfragen.
4. Definition abstrakter Begriffe.
5. Ungeordnete Worte zu einem Satz ordnen.
6. Bildbetrachtung: spontane Erklärung.
7. Ergänzung von Textlücken.
8. In einer Minute zu einem vorgesprochenen Wort 3 Reime finden.

[1]) Das Prüfmaterial wurde früher vom Institut für angewandte Psychologie in Kleinglienicke, jetzt wird es in dessen Auftrag von Dipl.-Ing. Buth, G. m. b. H., Charlottenburg, Königin-Luisenstr. 14, geliefert. Eine von Bobertag verfaßte Gebrauchsanweisung hat das genannte Institut herausgegeben.

Wie ein Blick über das Testverzeichnis lehrt, handelt es sich fast durchweg um Alternativtests, d. h. um solche, deren Lösung nicht abstufbar ist, sondern nur als richtig oder falsch (+ oder —) gewertet werden kann. Nichtbeantwortung wird ebenso wie eine offenkundig falsche Antwort als Minus gerechnet. Freilich ist es im Einzelfalle oft schwer genug, sich für das Prädikat + oder — zu entscheiden; hier kann nur eine lange Übung und eine gleichmäßige Innehaltung der einmal für die Entscheidung gewählten Kriterien Sicherheit gewähren.

Die Mannigfaltigkeit der Tests ist groß.

Das Gedächtnis wird geprüft einerseits in der Form der Merkfähigkeit für Ziffern und Sätze verschiedener Länge, für eine vorgelesene Geschichte und für drei zugleich aufgegebene kleine Aufträge, andererseits in der Form des Besitzstandes an elementaren Alltagskenntnissen (Wochentage, Monate, Münzen, rechts und links). Reichtum und Verfügbarkeit des Wortschatzes wird festgestellt durch die Anzahl der in drei Minuten zu nennenden Worte.

Seit 1911 ist auch ein Test für die Suggestibilität (bei Beurteilung von Linienlängen) eingefügt.

Motorische Fähigkeiten prüfen einige Tests des Abzeichnens, Ausschneidens, Schreibens. Mit praktischen Fertigkeiten hat man es bei dem Abzählen von Münzen, dem Herausgeben auf ein größeres Geldstück, dem Ausführen der oben genannten drei Aufträge zu tun.

Die meisten Tests gehen aber noch unmittelbarer auf eigentliche Intelligenzleistungen. Das Vergleichen bzw. Unterscheiden kommt in sehr verschiedenen Formen vor als: sensorisches Vergleichen (von gleich aussehenden Kästchen verschiedener Schwere), logisches Unterscheiden aus dem Gedächtnis, und zwar von Konkretem (Holz und Glas, Fliege und Schmetterling) und von Abstraktem (Lüge und Irrtum), ästhetisches Vergleichen (von abgebildeten schönen und häßlichen Gesichtern). Ferner gibt es Tests für das Definieren von Konkretem wie von Abstraktem; das Ergänzen von Lücken in einem Text; das Kombinieren eines Satzes aus drei Worten; das Ordnen, und zwar sowohl das sensorische Ordnen (von fünf Kästchen in einer Reihe nach ihrer Schwere), wie das sprachlich-logische Ordnen (der durcheinander gewürfelten Werte eines Satzes); das verstehende Auffassen eines Bildes; das kritische Auffassen, und zwar optisch (Bemerken von Lücken an gezeichneten Figuren) und logisch (das Erkennen von Widersinnigkeiten in bestimmten Sätzen); die praktisch-moralische I. (durch Fragen in der Form: Was würdest du tun, wenn das und das geschähe?).

Manche Tests wiederholen sich bei verschiedenen Altersstufen derart, daß dann jedesmal der Maßstab der Anforderung ein anderer ist. So werden die Bilder allen Jahrgängen vorgelegt; ein bloßes Aufzählen der abgebildeten Objekte entspricht dem I.-Stand der Dreijährigen, ein Beschreiben der Handlungen, welche die Personen ausführen, dem der Siebenjährigen, eine Erfassung des Zusammen-

hangs dem der Zwölfjährigen. Das Definieren konkreter Wörter kommt vor in den Staffeln der Sechsjährigen und der Neunjährigen; bei jenen genügt Definition durch bloße Zweckangabe (Was ist ein Pferd? „Zum Reiten"), bei diesen wird eine darüber hinausgehende verlangt (z. B. Was ist ein Pferd? „Ein Tier"). Endlich sind die Merkfähigkeitstests für Ziffern und Sätze nach der Länge vielfach abgestuft; so sollen die Dreijährigen zwei, die Vierjährigen drei, die Siebenjährigen fünf, die Zwölfjährigen sieben Ziffern nach einmaligem Hören nachsprechen können.

Terman (*118*) hat dieses Prinzip der Durchführung entsprechender Tests durch möglichst viele Altersklassen noch weiter ausgebildet, um die Vergleichbarkeit der Jahrgänge exakter zu gestalten: teils hat er die Tests selbst aufsteigend erschwert, teils hat er sie in gleichbleibender Form vorgelegt, aber steigende Anforderungen an die Art der Lösung gestellt.

Die Technik der BS.-Methode ist durchaus nicht so leicht, wie es bei der Einfachheit des Instrumentariums zunächst scheinen möchte. Es ist zu empfehlen, daß, soweit irgend angängig, der Prüfer stets mit einem Protokollführer zusammen arbeite, damit die schädliche Aufmerksamkeitsteilung auf Prüfen und Protokollieren vermieden werde. Beide Experimentatoren müssen sich eine große Übung erworben haben und gut aufeinander eingearbeitet sein, ehe sie zu ihren eigentlichen Versuchen schreiten. Der Prüfer muß eine fast mechanische Treffsicherheit und Gleichmäßigkeit im Formulieren der immer wiederkehrenden Fragen, im Tonfall usw. besitzen; er muß aber auf die vielen individuellen Variationen, die durch das verschiedene Verhalten der Prüflinge eintreten, gefaßt sein und bestimmte Maßnahmen dafür in Bereitschaft halten. Nie darf er durchblicken lassen, daß er mit manchen Antworten mehr, mit anderen weniger zufrieden ist, sondern muß stets die gleichmäßige ruhige Freundlichkeit bewahren. Der Protokollführer möge sich nicht nur auf die Aufzeichnung der Plus- und Minuszeichen beschränken, wie sie das Endergebnis jedes Tests ausdrücken, sondern auch in möglichstem Umfange die Aussagen des Prüflings sowie bemerkenswerte Verhaltungsweisen notieren. Dies ist einerseits nötig, weil die Entscheidung, ob Plus oder Minus, oft genug erst nachträglich bei ruhiger Überlegung möglich ist — wofür dann das Material vorliegen muß —, andererseits, weil ja auch eine qualitative Analyse des Prüflings ermöglicht werden soll.

Der einzelne Prüfling darf natürlich nicht nur mit den Tests seines Alters, sondern muß — wegen der bald zu besprechenden Staffelstreuung — mit einem beträchtlichen Bruchteil der ganzen Serie geprüft werden. Man beginne hierbei weder mit den allzu leichten noch mit den allzu schweren Tests, sorge für Abwechslung und füge, wenn Ermüdung merkbar sein sollte, kleine Pausen ein. Die individuelle Prüfung dauert je nach Alter und Umständen bei normalen Individuen 20—30 Minuten, bei nichtnormalen wegen der langsameren Reaktion $^1/_2$—$^3/_4$ Stunden.

Bei Massenuntersuchungen besteht die Schwierigkeit, daß schon Geprüfte und noch zu Prüfende miteinander in Verkehr treten können. Freilich ist die Gefahr einer Beeinflussung bei der Eigenart des Testmaterials nicht allzu groß; dennoch soll man, soweit es geht, die Möglichkeit einer Verständigung vermeiden. So empfiehlt es sich z. B. nicht, die Schüler einer Klasse an mehreren aufeinanderfolgenden Tagen zu prüfen.

3. Die Maßprinzipien der Staffelmethode.

Bei den I.-Maßen des Staffelsystems handelt es sich um „absolute Maße" in dem oben (S. 61) besprochenen Sinne.

a) Intelligenzalter (IA.). Indem wir auf Grund der Prüfung eines Kindes bestimmen, auf welcher Staffel die Aufgaben liegen, denen es zu entsprechen vermag, erhalten wir sein I.-Alter („niveau intellectuel", „mental age"). Ein Kind, das die für 9jährige bestimmten Tests zu lösen vermag, die für 10jährige aber nicht, hat das IA. 9 Jahr — ganz ohne Rücksicht darauf, ob sein Lebensalter 8 oder 9 oder 10 Jahr beträgt.

In Wirklichkeit besteht nun freilich niemals eine so scharfe, nach Teststaffeln bestimmte Grenze zwischen Können und Nichtkönnen. Denn die Tests innerhalb je einer Staffel sind nicht für alle Kinder von gleicher Schwierigkeit. Das eine Kind leistet a bis d, versagt aber bei e; ein anderes bewältigt a, c und e, aber nicht b und d. Dies liegt einerseits an den momentanen Schwankungen der Aufmerksamkeit, der Ermüdung usw., mit denen man natürlich immer rechnen muß, andererseits aber auch an qualitativen Unterschieden der I. Die Korrelation der einzelnen Seiten der I.-Funktionen ist eben nie eine so große, daß eine positive Leistung in a auch eine solche in den ungefähr „gleich schweren" Tests b bis d mit Notwendigkeit einschlösse.

Daraus ergibt sich, daß die Jahresstaffeln, die ein Kind voll beherrscht, und diejenigen, die zweifellos jenseits seiner Leistungs-

fähigkeit stehen, nicht scharf gegeneinander abgegrenzt sind, sondern daß es ein mehr oder minder breites Zwischengebiet gibt, in welchem positive und negative Leistungen unregelmäßig verstreut sind; wir wollen es das Gebiet der Staffelstreuung nennen.

Bei dem Herausziehen eines Mittelwertes aus diesem Gebiet der Staffelstreuungen ist Willkür nicht ganz zu vermeiden; doch ist die hierdurch eintretende Ungenauigkeit im allgemeinen belanglos. Das von Binet selbst vorgeschlagene Verfahren ist von seinen Nachfolgern gewöhnlich folgendermaßen abgeändert worden:

Jeder einzelne von den fünf Tests einer Staffel rechnet bei Lösung als ein fünftel Jahr. Hierbei können etwa nicht gelöste Tests tieferer Jahrgänge durch gelöste Tests höherer Jahrgänge wettgemacht werden[1]). So ergibt sich schließlich ein IA., das in ganzen und fünftel Jahren ausgedrückt ist.

Beispiel: Ein 8jähriges Kind hat die Tests seiner Altersstufe vollständig gelöst; aber von den Tests für 7jährige zwei verfehlt. Dagegen hat es noch von den Tests für 9jährige 3; für 10jährige einen gelöst. Von diesen 4 überzähligen Tests dienen 2 zur Kompensation der fehlenden 7-Jahr-Tests; die zwei übrigen sind als Zuschuß zum IA. 8 Jahr zu rechnen. Ergebnis:

$$\text{IA.} = 8\,{}^{2}/_{5}\text{ Jahr} = 8{,}4.$$

Schematisch läßt sich diese Berechnung so darstellen:

Alle Tests bis 6 Jahr einschließlich sind gelöst. Daher Ausgang der Berechnung:

6 Jahr,

Außerdem von Staffel 7 Jahr 3 Tests
" " " 8 " 5 "
" " " 9 " 3 "
" " " 10 " 1 "
zusammen 12 Tests = $2\,{}^{2}/_{5}$ Jahr,
IA. = 8,4 Jahr.

Die Staffelstreuung hat übrigens für die Berechnung des IA. eine methodologische Folge, die oft nicht genügend beachtet wird. Man muß, um etwaige Lücken aus niederen Teststufen auszugleichen, stets Tests aus höheren Jahrgängen zur Verfügung haben. Nun enthält z. B. die ursprüngliche Reihe von BS. nur Tests bis zur Stufe 13; deshalb ist es nicht möglich, das IA. 12 oder 13 korrekt zu berechnen, da möglicherweise Tests aus noch höheren Jahrgängen das Gesamtergebnis erhöht hätten. Man wird also bei Arbeiten nach BS. 1908 die Berechnungen nur bis IA. 11 führen dürfen.

---

[1]) Ein von mir früher gemachter Vorschlag, die Lösung der schwereren Tests höherer Jahrgänge auch mit einer höheren Wertziffer anzusetzen als die leichteren Tests niederer Jahrgänge, scheint technisch unbequem zu sein; er hat wenigstens bis jetzt noch keine Anwendung gefunden. Ich verzichte daher auf seine Wiederholung.

Die Staffelstreuung liefert noch einen anderen Maßwert als das IA., nämlich ihre eigene Streuungsbreite. Ein Kind, bei dem sich die Treffer und Versager unregelmäßig über die Tests der Altersstaffeln von 6 bis zu 10 Jahren verbreiten, hat zwar dasselbe IA., aber eine ganz andere Streuungsbreite als ein Kind, bei dem das Durcheinander von Treffern und Versagern nur in den Staffeln von 7—9 Jahr vorkommt.

Auf die Bedeutung, welche der verschiedenen Staffelstreuung zukommen kann, hat zuerst Bobertag (*51*) aufmerksam gemacht. Wir werden insbesondere bei schwachsinnigen Kindern abnorm große Streuungswerte finden.

Die Berechnung der Staffelstreuung kann einfach durch die „mittlere Abweichung" geschehen, indem man innerhalb des Streuungsgebiets für alle gelösten Tests die Abweichung vom IA. berechnet, aus diesen Werten ohne Beachtung des Vorzeichens das Mittel nimmt und dies zum IA. in Beziehung setzt.

Berechnung der Streuungsbreite für obiges Beispiel (IA. = 8,4 J.).

Abweichung	der	Staffel	7	Jahr	von	8,4 = 1,4	3 Tests	= 4,2	Jahr,
„	„	„	8	„	„	8,4 = 0,4	5 „	= 2,0	„
„	„	„	9	„	„	8.4 = 0,6	3 „	= 1,8	„
„	„	„	10	„	„	8,4 = 1,6	1 „	= 1,6	„
							12 Tests	= 9,6	Jahr.

$$\text{Mittl. Abweichung} = \frac{9{,}6}{12} = 0{,}8 \text{ Jahr.}$$

$$\text{Relative mittl. Abweichung} = \frac{\text{MAbw.}}{\text{IA.}} = \frac{0{,}8}{8{,}4} = 0{,}09.$$

b) Intelligenzrückstand und -vorsprung. Das IA. erhält erst einen Sinn, wenn es zum Lebensalter (LA.) in Beziehung gesetzt wird. Diese Beziehung wurde von Binet und vielen seiner Nachfolger als einfacher Differenzwert IA.—LA. behandelt, der die Spannung zwischen vorhandener und geforderter (als Norm zu erwartender) I. ergibt. Ist dieser Differenzwert negativ, so bedeutet er Intelligenzrückstand (IR.), ist er positiv, Intelligenzvorsprung (IV.). IR. = — 2 bedeutet also: das Kind ist in seiner geistigen Entwicklung um 2 Jahre hinter dem normalen Stand seines Alters zurück. Es steht z. B. bei einem Lebensalter von 8 Jahren auf der I.-Höhe der 6jährigen.

Binet und viele spätere Anwender der Staffelmethode haben gemeint, damit ein völlig ausreichendes Stufenmaß der I. gewonnen zu haben. In der Tat ist es auch für manche Zwecke, namentlich bei Vergleichung gleichaltriger Kinder ausreichend; aber der Fort-

schritt der Forschung ließ doch entschiedene Mängel dieses Maßes hervortreten. Der eine Mangel: die zu grobe Ganzzahligkeit des Maßwertes, läßt sich leicht abstellen. Von Binet waren sowohl LA. wie IA. lediglich als ganze Zahlen in Rechnung gestellt worden; dadurch kamen große Ungenauigkeiten in die Bestimmung.

Ein Kind, das zwei Tests über die Staffel 8 Jahr gelöst hatte, bekam dennoch nur IA. 8. War das Kind 8 Jahr 8 Monate alt, so bekam es das LA. 9; und es ergab sich nach der Formel IA. — LA. ein Rückstand von einem Jahr, obgleich in Wirklichkeit nur ein ganz unbedeutender Rückstand vorliegt. Die genauere Berechnung ergibt: IA. = 8,4; LA. = 8,67; IR. = — 0,27.

Wichtiger ist der zweite Mangel. Binet hatte in dem Wert wohl eine Konstante gesehen; er betrachtet einen IR. von 2 Jahren generell als Ausdruck einer bestimmten, und zwar schweren Rückständigkeit, ohne dabei weiter auf das Alter Bezug zu nehmen. Auch von seinen Nachfolgern sind oft genug gleiche I.-Rückstände als gleichwertig betrachtet und ohne Rücksicht auf das Lebensalter zu statistischer Verrechnung vereinigt worden. Das ist aber unzulässig, denn der Wert IR. ist lediglich ein Momentanwert. Ein Kind, das mit 8 Jahren auf dem IA. 6 steht, hat nicht etwa mit 12 Jahren das IA. 10, sondern ein noch tieferes. IR. bleibt also nicht konstant; die Bezeichnungen „um 1 (2) Jahr rückständig", „um 1 (2) Jahr voraus" haben nur so lange eindeutigen Sinn, als sie auf ein bestimmtes Lebensalter bezogen werden.

c) Der Intelligenzquotient. Um die Vieldeutigkeit des Differenzwertes zu beseitigen, schlug ich (*113*) 1912 die Anwendung des Intelligenzquotienten $\left(\text{IQ.} = \frac{\text{IA.}}{\text{LA.}}\right)$ vor. Dieser drückt aus, welchen Bruchteil der Durchschnittsintelligenz ein Kind besitze. Für ein genau durchschnittliches Kind muß dieser Wert unabhängig vom Lebensalter = 1 sein, bei den Kindern mit I.-Rückstand ist IQ. ein echter Bruch, bei denen mit I.-Vorsprung ein unechter Bruch.

Die Hoffnung, mit dem IQ. eine wirkliche persönliche Konstante, also einen Dauerwert der I. gewonnen zu haben, hat sich innerhalb gewisser Grenzen erfüllt. Dies wird durch die Studien von Weigl (*122, 174*), Descoeudres (*58*), Kuhlmann (*77, 78*) an geistig schwachen Kindern und vor allem durch die Untersuchungen von Terman (*119, 171*) an normalen Kindern erwiesen. Konstant kann der IQ. nur dann sein, falls die I.-Entwicklung der Kinder bei verschiedenem absoluten Tempo doch im großen und

ganzen proportional zur normalen Entwicklung verläuft, derart, daß das prozentuale Verhältnis in den verschiedenen Altersstufen gleichbleibt. (Wenn z. B. ein Kind mit 8 Jahren $^3/_4$ der normalen I., also IA. 6 besitzt, müßte es auch mit 12 Jahren $^3/_4$ der Normal-I., also IA. 9 besitzen. Beidemal wäre IQ. = 0,75.) Dieser proportionale Verlauf kann natürlich, wenn überhaupt, so nur in den Zeiten wirklicher I.-Entwicklung bestehen; denn wenn ein schwachsinniges Kind in einem Alter schon zum I.-Stillstand kommt, in welchem sich das normale Kind noch immer weiter entwickelt, dann muß das Verhältnis zwischen IA. und LA. mit steigenden Jahren immer schlechter werden. Nun wurde aber auch für die Zeit der I.-Entwicklung selbst die Konstanz des IQ. von Bobertag bezweifelt; er glaubte, daß der IQ. auch schon bis zum 12. Lebensjahre langsam aber stetig abnehme. Die neueren Untersuchungen haben diese Vermutung nicht bestätigt. Sie zeigen vielmehr, daß schon die gegenwärtigen Staffelmethoden für die Altersspanne zwischen 6 und 12 Jahren die Feststellung annähernd konstanter IQ. erlauben, die einerseits als Kennwerte für den I.-Habitus einzelner Kinder und bestimmter Schwachsinns- und Hochbegabungsgrade, andererseits als Maßwerte der I.-Verteilung und der vergleichenden Charakteristik verschiedenen Gruppen dienen können. Vielleicht wird eine Verbesserung und Erweiterung der Staffelmethoden erlauben, die Messung des IQ. auch noch über das vollendete 12. Lebensjahr hinauszuführen. Bei Terman liegen bereits Ansätze dazu vor.

Bei der Berechnung des IQ. müssen IA. und LA. mit ersten Dezimalstellen bestimmt werden; dann kann man den IQ. bis zur zweiten Dezimale berechnen.

4. Mängel der Staffelmethode und Änderungsvorschläge.

Das BS.-System hat auf der einen Seite blinde und kritiklose Anhängerschaft, auf der anderen heftige Kritik, ja völlige Ablehnung erfahren; es wird unsere Aufgabe sein, zwischen diesen beiden Einseitigkeiten die rechte Mitte zu halten.

Die Kritik kann sich auf dreierlei beziehen: auf das Prinzip des Verfahrens, auf die Tragweite, die man ihm beimißt, auf die gegenwärtige Form der Technik und Berechnung.

Grundsätzlich muß die Methode natürlich von allen denen abgelehnt werden, welche das Dasein einer I. als einer einheit-

lichen Beschaffenheit des Individuums von vornherein leugnen. Wenn man — wie Ziehen, Höper und andere (s. S. 9) — der Meinung ist, daß die einzelnen bei einer solchen Prüfung getroffenen Funktionen als verschiedene Seiten der Persönlichkeit unvermittelt nebeneinander bestehen, ohne daß ein in ihnen allen hervortretender Grad geistiger Regsamkeit und Anpassungsfähigkeit sie zusammenhielte, dann kann man auch nicht zulassen, daß die Einzelergebnisse auf einen Generalnenner, den allgemeinen „I.-Grad" dieses Individuums gebracht werden.

Wir können hier auf eine Widerlegung dieses Einwandes verzichten, da an anderen Stellen des vorliegenden Buches dargelegt wurde, aus welchen Gründen und in welchem Sinne wir die I. als eine einheitliche Funktion ansehen (vgl. z. B. S. 28). Sie ist für uns nicht eine feste Zwangsjacke, die alle einzelnen Leistungsgebiete in eine eindeutige gemeinsame Form preßt, sondern sie drückt nur die Fähigkeit aus, sich auf geistige Ziele mit einer größeren oder geringeren Leichtigkeit einzustellen, wobei der relative Anteil der einzelnen Funktionen noch ziemlich variieren kann. Demnach darf das aus der Staffelprüfung zu errechnende IA. (bzw. die aus ihm abgeleiteten Maßwerte) nicht als eine völlig eindeutige Festlegung der I. eines Prüflings betrachtet werden, sondern nur als eine ganz grobe Charakteristik des quantitativen I.-Wertes, unbeschadet der qualitativen Unterschiede; denn ein und dasselbe IA. kann aus Staffelstreuungen der verschiedensten Art abgezogen werden. Aber hierin scheint nicht ein Nachteil, sondern ein Vorteil des Begriffs des IA. zu liegen. Denn er wird dadurch zum Ausdruck einer schon früher erwähnten psychologischen Grundtatsache: bei dem nur formalen Charakter der I. und der nicht vollständigen Korrelation der in ihr enthaltenen Teilfähigkeiten gibt es niemals wirkliche phänomenologische Gleichartigkeit der I. bei zwei Individuen. Wohl aber gibt es — im Hinblick auf die einheitliche Zielsetzung aller I., nämlich die Anpassung an neue Anforderungen — eine teleologische Gleichwertigkeit. Und für diese Gleichwertigkeit trotz fehlender Gleichartigkeit soll das IA. ein Annäherungsmaß darstellen.

Das Recht, in diesem eingeschränkten Sinne von einem einheitlichen Maß der I. eines Menschen zu sprechen, kann aber nun auch a posteriori aus den Ergebnissen abgeleitet werden. Wäre jener Maßwert nur das Summationsergebnis von beliebigen Einzelwerten, die eigentlich nichts miteinander zu tun haben, dann wäre

es unbegreiflich, wie er diagnostisch und massenstatistisch zu unleugbaren Befunden von praktischer und wissenschaftlicher Bedeutung hätte führen können[1]). Ja noch mehr. Hätten die geprüften Einzelfunktionen keine bedeutende Korrelation untereinander, dann wäre es unwahrscheinlich, daß sie in gewissen Individuen überwiegend mit geringen Leistungsgraden, in anderen überwiegend mit hohen Leistungsgraden auftreten. Es müßten vielmehr in allen Individuen hohe, mittlere und niedere Leistungsgrade beliebig gemischt vorhanden sein und daher die Resultantenwerte alle sehr ähnlich sein. In Wirklichkeit aber zeigt die Verteilung der Resultantenwerte eine solche relative Häufigkeit von guten, mittleren und schwachen Werten, wie sie nur bei merklicher Korrelation der beteiligten Einzelfunktionen denkbar sind.

Der im IA. (und seinen Ableitungen) zu gewinnende Maßwert ist also zulässig und brauchbar; aber er ist in keiner Weise erschöpfend; er verlangt nach Ergänzung, damit das Bild der geprüften I. vervollständigt werde. Solche Ergänzungen liefert zum Teil die Staffelmethode selber. Zunächst bietet sie ja noch einen zweiten Maßwert, die „Staffelstreuung", dar; dieser gibt an, ob die I. des Prüflings sich einigermaßen harmonisch auf gleichmäßigen Leistungshöhen der einzelnen Teilfunktionen bewegt, oder ob sie disharmonisch auffallend hohe mit auffallend niederen Leistungen verbindet. Aber die qualitativen Unterschiede der Streuung werden auch hiermit nicht erreicht, und deshalb wird es in vielen Fällen nötig sein, außer den beiden Resultantenwerten (IA. und Staffelstreuung) auch eine eingehendere Analyse der Prüfung zu geben. Oft wird es gerade von Wichtigkeit sein, festzustellen, in welchen Tests eine besondere Rückständigkeit, in welchen ein besonderer Vorsprung vorhanden war. Auch die Beobachtung des Kindes während der Prüfung darf man nicht gering schätzen, denn bei manchen Tests gibt es Verhaltungsweisen, die von hohem Interesse sein können (z. B. auch für die medizinische oder pädagogische Beurteilung), ohne daß dies in dem Plus oder Minus des Testergebnisses zum Ausdruck käme. Es sei hier u. a. auf die Art der Bildbeschreibung und auf das Aufzählen von 60 Worten hingewiesen, sodann auf das Verhalten des Kindes bei dem Ordnen der fünf gleichaussehenden, aber verschieden schweren Kästchen. Ob hier schließlich die richtige Ord-

[1]) Über den Wert und die Tragweite dieser Befunde wird im dritten Hauptteil dieses Buches berichtet.

nung getroffen wird, ist lange nicht so wichtig wie die Beobachtung des kindlichen Benehmens bei der Arbeit: ob und wie schnell es die ungewohnte Aufgabe erfaßt, ob es nur immer zwei Gewichte vergleicht oder jedes Gewicht beim Einordnen mit allen anderen vergleicht usw.[1]). So ist also bei diesen Untersuchungen vor einem bloßen Jagen nach Zahlenwerten zu warnen; Rechnung und qualitative Analyse müssen sich ergänzen, wobei natürlich je nach der Fragestellung die eine oder andere Methode bald mehr, bald weniger hervortreten wird.

Neben diesen der Staffelmethode selbst zu entnehmenden quantitativen und qualitativen Angaben wird man stets, wo immer es möglich ist, nach weiteren Vervollständigungen des I.-Bildes suchen müssen, wie sie andere Prüfungsmethoden und vor allem die durch kein Experiment jemals ersetzbare Beobachtung gewähren. Die Staffelmethode an sich ist nur das „psychographische Minimum", das eine kurze und schnelle Vororientierung und eine bequeme Vergleichung gestattet: praktische Entscheidungen über Schulwahl, Berufsschicksal und ähnliches sollte man aber niemals allein auf eine solche Prüfung zu stützen wagen.

Erkennt man das Prinzip der Staffelmethode als berechtigt an, so bietet doch die Methode in ihren Einzelheiten vielfachen Anlaß zu berechtigten Bemängelungen und zu Reformvorschlägen. Schon die erste Bekanntschaft mit den Tests läßt ihre Ungleichwertigkeit erkennen. Manche sind recht gut, wie das Definieren, die Bildbeschreibung, die Vergleiche, die I.-Fragen. Wertvoll ist es ferner, daß auch „stumme" Tests vorhanden sind, so der ästhetische Vergleich, das Erkennen von Lücken auf Bildern und vor allem der treffliche Test des Ordnens der 5 Gewichte; aber immerhin haben schon diejenigen Kritiker Recht, welche das Überwiegen von Tests, die auch sprachliche Gewandtheit verlangen, und die Vernachlässigung von „Handlungs"-Tests bemängeln. Auch die Formulierung einiger Tests ist ungeschickt oder ungeeignet; so wirkt die Blutrünstigkeit der von Binet vorgelegten „Absurditäten" abstoßend. Im ganzen merkt man dem System bis heute eine gewisse Grundsatzlosigkeit an; die Zusammenstellung war nach den damaligen recht begrenzten Erfahrungen und den zufälligen Inter-

[1]) Vgl. zu dem Test S. 121/2.

essenrichtungen der Schöpfer erfolgt, und auch die vereinzelten Abänderungen, die von Bobertag, Terman u. a. vorgenommen worden sind, haben an diesem Eindruck nichts Entscheidendes ändern können. Wichtige Seiten der I. sind überhaupt nicht berührt, andere, die nur sekundäre Bedeutung haben (z. B. die Merkfähigkeit), sehr ausführlich berücksichtigt. Manche Tests sind in hohem Grade von äußeren, namentlich häuslichen Einflüssen abhängig (Hersagen der Wochentage); selbst solche, welche direkt vom Schulwissen beeinflußt sind, fehlen nicht (Diktat, Rückwärtszählen usw.), obgleich die Ausschaltung bloßer Schulkenntnisse doch eine der ersten Aufgaben einer Fähigkeitsprüfung sein sollte.

Die Zähigkeit, mit der man an dem einmal aufgestellten System festhielt, ist allerdings aus einem technischen Grunde verständlich; eine Änderung in der Testauswahl bietet beim Staffelsystem nämlich eine ganz andere Schwierigkeit, als wenn es sich um eine lose Testreihe handelte. Denn da jeder Test an der Eichung der einzelnen Staffeln beteiligt ist, kann eine Änderung möglicherweise das Gleichgewicht des Systems und die Richtigkeit aller Berechnungen erschüttern.

Die bisher vorgenommenen Verbesserungen des Systems beziehen sich vor allem auf eine Zurechtrückung der Tests, d. h. auf ihre anderweitige Zuordnung zu bestimmten Altersstufen. Binets Anordnung enthielt hier offenkundige Unstimmigkeiten: die Tests für die jüngeren Kinder waren zu leicht, die für die älteren zu schwer. Die Nachprüfungen von Bobertag, Terman, Jaederholm an größeren Prüflingsmassen ergaben folgende Gesichtspunkte für die Verschiebung der Tests:

1. Ein einzelner Test ist derjenigen Altersstufe zuzuordnen, auf der er nach steilem Altersfortschritt von etwa 75% der Prüflinge gelöst wird (vgl. S. 168). Es mögen z. B. für zwei Tests die Lösungsprozente betragen:

	Alter:	8 Jahr	9 Jahr	10 Jahr
Prozentzahl der Prüflinge, welche den Test lösen.	Test I:	30%	40%	75%
	Test II:	60%	65%	75%

Dann ist trotz der gleichen Lösungsprozente, die beide Tests für das Alter 10 Jahr aufweisen, Test I für dieses Alter viel charakteristischer, weil zu ihm ein steiler Altersfortschritt hinführt.

2. Die Zusammenstellung von Tests für die einzelnen Altersstufen ist dann angemessen, wenn sich auf jeder Altersstufe eine symmetrische Verteilung der IA. ergibt; d. h. es müssen auf jeder

Altersstufe diejenigen, die auf der I.-Höhe des Alters stehen, die kompakte mittlere Masse bilden; auf ihren beiden Seiten müssen sich zu ungefähr gleichen Bruchteilen diejenigen gruppieren, die unter dem Niveau und die über dem Niveau stehen. Es muß also das durchschnittliche IA. der Gruppe dem LA. entsprechen. Im Verfolg dieses Gedankens haben Terman und Jaederholm die Tests von Binet so lange über die verschiedenen Altersstufen hin und her geschoben, bis der Durchschnitt der Massenprüfung auf allen Altersstufen der Gleichung IA. = LA. genügte.

In Zukunft aber werden noch weitere Verbesserungen nötig sein, die sich nicht nur auf Verschiebung der Tests, sondern auf die Umgestaltung des Testbestandes selber erstrecken. Die Bedürfnisse, die hier befriedigt werden müssen, sind mannigfacher Art. Zunächst müssen die zweifellos ungeeigneten Tests beseitigt und durch geeignetere ersetzt werden.

Ferner machte sich bisher das Fehlen von Teststaffeln für die höheren Altersstufen sehr unangenehm bemerkbar. Was von Binet selbst an Tests über den Jahrgang 11 hinaus geschaffen war, hatte noch ganz den Charakter des Tastenden und Vorläufigen; die Tests konnten höchstens dazu dienen, die nötige Staffelstreuung zur korrekten Bestimmung der IA. 10 und 11 zu liefern, aber zu einer selbständigen Feststellung für die IA. 12—15 reichten sie nicht aus. Das neue System von Terman (*118*) bringt bereits Staffeln für die Altersstufen 14, 16 und 18 Jahr; aber diese werden wohl auch noch nichts Endgültiges bedeuten. Sie sind nämlich meist nach dem Grundsatz hergestellt, daß die Aufgaben denen früherer Jahrgänge der Art nach gleichen, nur in der Schwierigkeit gesteigert sein sollen. Nun ist aber die I. der höheren Jugendjahre nicht so sehr durch eine Gradsteigerung gegenüber den vorangegangenen Altersstufen, als durch qualitative Wandlungen gekennzeichnet (s. S. 33); und deshalb müssen die hier zu stellenden Aufgaben einem anderen Typ angehören als beim kleineren Kinde. Es ist zu hoffen, daß die neue Wendung der Testforschung, wie sie durch die Begabungsauslese herbeigeführt worden ist, solche Aufgaben liefert und damit auch eine Erweiterung des Staffelsystems möglich macht. Unsere eben erscheinende Methodensammlung (*8*, III) dürfte bereits manches geeignete Material für höhere Altersstaffeln enthalten.

Sehr wünschenswert ist ferner eine Erweiterung des Systems durch Schaffung von Parallelreihen für alle Jahrgänge. Wie

gern würden wir die Methode dazu benutzen, die I.-Entwicklung derselben Kinder durch mehrere Jahre hindurch zu verfolgen; das macht aber Schwierigkeiten, weil ja das Kind den gleichen Tests, denen es schon früher ausgesetzt gewesen, im Wiederholungsfall nicht mehr ebenso gegenübersteht. Hätten wir dagegen andere, aber gleichwertige Serien, so wären mehrfach wiederholte Prüfungen derselben Kinder möglich. Desgleichen könnten bei Massenuntersuchungen solche Kinder, zwischen denen die Gefahr der Aussprache besteht, mit verschiedenen Serien geprüft werden. Endlich ist es wertvoll, für den Fall, daß eine Untersuchung durch Störung oder Ungeschick unbrauchbar wird, über ein Ersatzverfahren zu verfügen.

Der Aufgabe, Paralleltests zu schaffen, ist bisher nur Karstädt (*73*) nähergetreten. Er stellte zunächst ihre Notwendigkeit fest; es zeigte sich nämlich, daß bei Anwendung der gewöhnlichen Reihe die späteren Prüflinge einiges über die zu erwartenden Aufgaben von den früheren Prüflingen gehört hatten, ferner daß bei Wiederholung der Prüfung an denselben Kindern selbst nach Monaten noch deutliche Erinnerungsspuren vorhanden waren. Sodann versucht er, eine Reihe von neuen Tests zu eichen, die als Ersatz- und Ergänzungstests benutzt werden können. Seine Eichung bezieht sich auf folgende Testarten: Definition mit Zweckangaben und mit Oberbegriffen, leichte und schwere Verstandesfragen, Vergleiche aus der Erinnerung, Datum, Dreiwortmethode, Kritik von absurden Sätzen. Nach den oben angegebenen Grundsätzen der Eichung (etwa 75% Lösungen, steiler Altersfortschritt) wählte er die Tests aus, die für bestimmte Altersstufen geeignet erschienen. Indessen können diese Feststellungen noch nicht als endgültige betrachtet werden; dazu ist die Zahl der Prüflinge viel zu klein gewesen. Auch bediente sich K. des schriftlichen Massenverfahrens; es ist aber nicht ohne weiteres zulässig, die hierbei gewonnenen Ergebnisse auf das mündliche individualisierende Verfahren zu übertragen, wie es bei der BS.-Prüfung angewandt werden muß. Immerhin sind K.s Bemühungen als schätzenswerte Vorarbeiten zu künftiger Herstellung von Parallelreihen zu betrachten.

Eine besondere Aufgabe besteht in der Schaffung von Staffelsystemen für mindersinnige Kinder. Bei Blinden müßten alle optischen Tests, bei Taubstummen alle sprachlichen Tests durch andere ersetzt werden. Auch hier ist das Bedürfnis nach einem

exakten Maßstabe der I. lebhaft; denn die Frage, ob die Leistungsausfälle des Kindes lediglich auf dem Fehlen des einen Sinnes beruhen oder ob sie durch eine damit verbundene geistige Minderwertigkeit mitbedingt sind, ist sonst nur schwer zu beantworten. Dies ist besonders brennend bei Taubstummen, deren Defekt ja häufig zugleich ein solcher des Geisteszustandes ist. „Namentlich die Worttauben, die zwar hören, wenn man sie etwas fragt, die aber das Gehörte nicht begreifen, werden nicht selten, völlig zu Unrecht, für erziehungsunfähige Idioten gehalten[1])." Und so ist denn hier auch bereits der Versuch gemacht worden, eine eigene Teststaffel zusammenzustellen und zu erproben; sie stammt von D. Herderschêe (*68*) in Amsterdam, der sie an 82 Insassen der dortigen Taubstummenanstalten anwandte.

Die Liste (siehe auch MS. 243 ff.) besteht natürlich durchweg aus „stummen Tests", die sich unter anderem auf folgende Aufgaben beziehen: Erkennen von Gegenständen durch den Tastsinn, Sortieren von größeren Gegenstandsmengen, Ordnen abgestufter Gegenstände nach Größe, Schwere, Form, Wiederfinden eines vorher gezeigten Gegenstandes in einer größeren Menge, Erkennen von Bildabsurditäten, Zusammensetzen von Figuren und ähnliches. H. sieht selbst in seiner Serie nur einen vorläufigen Vorschlag; sowohl die Testauswahl wie auch die Alterseichung der einzelnen Tests bedarf noch gründlicher Nachprüfung. Eine solche ist in Hamburg bereits in Angriff genommen.

Es gibt noch einige weitere Vorschläge von Prüfungssystemen, die sich auf den BS.schen Grundlagen aufbauen; aber sie enthalten Änderungen grundsätzlicher Art. Bald wird zwar das Staffelprinzip beibehalten, aber neben der Jahresgruppierung noch eine andere Einteilung der Tests versucht (Meumann [*9*]); bald werden zwar die alten Tests im wesentlichen beibehalten, aber eine neue Art der I.-Berechnung unter Verzicht auf das IA. vorgeschlagen (Yerkes [*125*], Schultze [*107*], Treves und Saffiotti [*103*, *104*])[2]).

5. Begabungs- und Entwicklungstests. (Meumanns Vorschlag.)

Meumann (*9*, 771 ff.) hat kurz vor seinem Tode in der zweiten Auflage seiner „Vorlesungen" den Entwurf eines neuen Staffelsystems veröffentlicht, das sich möglichst nahe an die Binet-

[1]) Herderschêe (*68*, S. 40).

[2]) Einige Änderungsvorschläge, die in diesem Buch nicht im einzelnen behandelt werden, findet man besprochen bei Meumann (*9*) und in der Methodensammlung (8, III).

Simon-Tests anlehnt, aber für jede Altersstufe 10 Tests aufstellt. Bedauerlicherweise ist der Vorschlag praktisch unbrauchbar.

Meumann gibt an, daß das System aus zahlreichen Vorbesprechungen mit Lehrern und vielen Vorversuchen hervorgegangen sei; leider aber tritt das Ergebnis dieser Vorarbeiten nirgends zutage. Vielmehr wird die Liste ohne Anweisung über Technik und Berechnungsweise, ohne Beleg für die Gründe der Testzuweisung zu einer bestimmten Altersstufe oder zu einer bestimmten Testgruppe veröffentlicht; und der ratlose Leser hat den Eindruck, daß hier eine noch unabgeschlossene Arbeit verfrüht dem Publikum zur Verfügung gestellt worden ist.[1]) Manche Tests wiederholen sich (nach dem Vorbild von Terman) in mehreren Altersstufen; aber man vermißt eine Angabe darüber, welche verschiedenen Anforderungen an deren Lösung bei jüngeren und älteren Kindern gestellt werden. Eine Neuerung Meumanns besteht darin, daß er innerhalb jeder Altersstufe eine Dreiteilung der Tests vornimmt; der an und für sich wertvolle Gedanke nämlich, daß die drei Hauptbedingungen der Leistung: Entwicklung, Begabung und Umwelt bei den verschiedenen Tests in sehr verschiedenem Maße beteiligt sind, wird ohne weiteres zu einer Einteilung der Tests in Entwicklungs-, Begabungs- und Umwelttests benutzt; und zwar wird diese Einteilung nicht auf Grund eigens dazu vorgenommener Untersuchungen gegeben, sondern a priori und mit einer solchen Willkür, daß manche Tests für die eine Altersstufe zu den Entwicklungs-, für die andere zu den Begabungstests gerechnet werden[2]). Bezüglich alles Näheren mag auf die eingehende Kritik Hyllas verwiesen werden.

Besondere Erörterung verdient aber die Frage, ob der neue Grundgedanke Meumanns, die Tests nach ihrer Abhängigkeit von Entwicklung, Begabung und Umgebung zu sondern, prinzipiell in einem Staffelsystem durchführbar ist.

---

[1]) In dem kurze Zeit später erschienenen Abriß der exp. Pädagogik (*11*) hat denn Meumann auch auf den Wiederabdruck seiner Skala verzichtet, „da die Mängel der ganzen Methode noch so große sind, daß alle Einzelheiten einer kritischen Nachprüfung bedürfen".

[2]) So gibt Hylla (*69*) über die Zuweisung des Tests „Nachsprechen von Ziffern und Sätzen" folgenden Auszug aus Meumanns Tabelle: „Dieser Test gilt für das 5. Lebensjahr vorwiegend als Begabungs-, in zweiter Linie als Entwicklungstest, für das 6. Jahr nur als Begabungstest, für das 7. vorwiegend als Entwicklungs- und nur nebenher als Begabungstest, für das 8. als reine Entwicklungsprobe; für das 9. dagegen soll er wieder hauptsächlich Begabungstest, weniger Entwicklungstest sein, während er im 10. und 11. Jahre als bloßer Entwicklungstest, im 12. außerdem als Begabungstest, im 13. Jahr gar nicht und im 14. und 15. Jahr ebenfalls in erster Linie als Entwicklungsprobe, in zweiter als Begabungstest registriert wird!" — An einer anderen Stelle (*9*, S. 286/7) wird der bekannte Test: zu einem Reizwort den über- oder untergeordneten Begriff oder den Ausdruck der Ursache oder der Wirkung zu suchen — als Begabungsprobe, dagegen der ganz analoge Test: zu einem Wort den gegensätzlichen Begriff zu finden — als Entwicklungsprobe bezeichnet.

Zunächst scheinen mir „Umgebungstests" überhaupt nicht in das System einer IP. hineinzugehören. Denn das System soll nicht eine Inventaraufnahme des tatsächlichen geistigen Bestandes, nicht die Prüfung der von außen her (durch Umgebungseinflüsse) bedingten Kenntnisse, sondern die Feststellung des Fähigkeitsgrades herbeiführen; und alle Bemühungen der Psychologen in den letzten Jahrzehnten gingen dahin, hierbei solche Aufgaben möglichst auszuschalten, die durch äußere Einflüsse (häusliche Umgebung, Schulunterricht usw.) bedingt sind. Diese Ausscheidung ist noch nicht voll erreicht; aber sie muß immer mehr angestrebt werden. Deshalb erscheint die ausdrückliche Hineinnahme von Umgebungstests in das System geradezu als ein Rückfall in frühere unvollkommenere Stadien unserer Arbeit[1]).

Viel bedeutsamer ist M.s Versuch, die beiden inneren Bedingungen der I.: Entwicklung und Begabung, zu trennen. Mit vollem Recht betont er (S. 106 und 284), daß manche psychischen Betätigungsgebiete ganz vorwiegend in Korrelation zum Lebensalter stehen; die Leistungen variieren stark mit den Jahren, während sie sich bei gleichaltrigen, aber verschieden begabten Kindern weniger deutlich unterscheiden; Meumann nennt als Beispiel die Spanne des unmittelbaren Behaltens. Die hierher gehörigen Tests wären also vornehmlich „Entwicklungstests". Andere Funktionsgebiete wiederum sind so sehr von der angeborenen Begabungshöhe abhängig, daß die Altersunterschiede weit zurücktreten; wir finden dann gute Leistungen bei begabten Kindern sehr verschiedenen Alters, ebenso verteilen sich die schlechten Leistungen über die unbegabten Kinder vieler Jahrgänge; man braucht ja hier nur an die Fähigkeit des Zeichnens zu erinnern. Der Prüfung solcher Funktionen dienen die „Begabungstests". Meumann versäumt freilich auch nicht, darauf hinzuweisen, daß für viele Gebiete beide Scheidungen zusammenfallen, d. h. daß das Kind geringerer Begabung sich ähnlich verhält wie das Kind niederen Alters, oder anders ausgedrückt, daß die geringere Begabung den Eindruck einer gewissen „Zurückgebliebenheit" macht (S. 107); auch beim Erwachsenen äußert sich dann die schwächere Begabung in einer Annäherung an die mehr kindlichen Leistungsformen („Infantilismus"). Gerade die eigentlichen Denkfunktionen gehören dieser Mischgruppe an.

[1]) Hierzu vergleiche man, was an früherer Stelle (S. 72 f.) über Kenntnisprüfungen gesagt worden ist.

Es ist nun ohne weiteres klar, daß eine psychographische Gesamtaufnahme eines kindlichen Intellekts, wenn eine solche überhaupt durch Tests möglich ist, jede dieser Funktionsgruppen prüfen muß; und als Ideal ist eine solche „Methode der intellektuellen Gesamtproben" gewiß anzustreben. Nur wäre diese etwas völlig anderes als die BS.sche Staffelmethode, und man darf nicht glauben (wie es Meumann tat), daß sich innerhalb dieser Methode durch Aufstellung getrennter Entwicklungs- und Begabungstests[1]) eine Umgestaltung herbeiführen lasse. Sonderbarerweise hat Meumann dort, wo er die Anwendung seiner Theorie auf das BS.-System zu machen sucht (S. 287/8), jene Mischgruppe von Tests mit zusammenfallender Alters- und Begabungssymptomatik unerwähnt gelassen; es ist jetzt nur noch von „Begabungs"- und „Entwicklungs"-Proben die Rede, und jeder Test gilt ihm zum mindesten vorwiegend als zur einen oder anderen Gruppe gehörig. In Wirklichkeit aber ist jene Mischgruppe nicht nur die stärkste, sondern auch die einzige, die der Grundidee der Staffelprüfung entspricht. Denn das war ja gerade der fruchtbare Gedanke Binets, daß der Begabungsgrad eines Kindes durch Vergleich mit der normalen Altersentwicklung festgestellt werden soll. Demnach sind hier vor allem solche Tests zu fordern, die mit dem Alter deutlich variieren und zugleich innerhalb jeder Altersstufe die Begabungsgrade der Kinder hervortreten lassen, also „Begabungs-Alters-Tests", wie man im Unterschied zu Meumanns Bezeichnungen sagen könnte.

Unsere künftige Aufgabe für den Ausbau der Staffelprüfung wird also meines Erachtens nicht darin bestehen, Entwicklungstests einerseits, Begabungstests andererseits auf jeder Staffel zusammenzustellen, sondern im Gegenteil einseitige (von der Begabung unabhängige) Entwicklungstests ebenso wie einseitige (vom Alter unabhängige) Begabungstests auszumerzen und möglichst nur Begabungs-Alters-Tests zu verwerten. Da zu diesen, wie Meumann selbst an der früheren Stelle hervorhob, gerade die spezifischen Denkproben gehören, so befinden wir uns damit recht eigentlich im Gebiet der I.-Tätigkeit im engeren Sinne.

Freilich wird uns hiermit von neuem klargemacht, daß die Staffelmethode eine durchaus begrenzte Bedeutung hat. Sie gibt kein Gesamtbild der geistigen Beschaffenheit eines Individuums, sondern nur eine Kennzeichnung derjenigen I.-Funktionen, deren

[1]) Von den Umgebungstests ist später zu sprechen.

Hoch- bzw. Tiefstand zugleich als beschleunigte bzw. verlangsamte Entwicklung angesehen werden kann. Die Anwendbarkeit der BS.-Methode und insbesondere die Konstanz des IQ. gewinnt so auch eine interessante theoretische Bedeutung: es ist nämlich dadurch erwiesen, daß in gewisser Hinsicht ein Recht dazu besteht, die schwache I. unter dem Bilde der „Zurückgebliebenheit" zu betrachten. Aber doch nur in gewisser Hinsicht. Die geringe Begabung ist, verglichen mit der normalen, nicht nur infantiler, sondern auch qualitativ andersartig; sie hat ihre typischen Züge, die nicht mit dem Alter korrelieren. Diese aber sind durch die Staffelmethode ihrem Grundprinzip nach nicht zu fassen. Demnach sind die von Meumann vorgeschlagenen Entwicklungs- und Begabungsproben wohl zur Vervollständigung des psychographischen Bildes zu fordern, aber außerhalb des Rahmens der in ihren Zielen begrenzten Staffelmethode.

Fragen wir schließlich, auf welche Weise sich die vorwiegende Zugehörigkeit bestimmter Tests zu einer der zwei Gruppen feststellen lasse. Hierzu sind vergleichende Untersuchungen notwendig über die Leistungen verschiedener Altersstufen einerseits, verschiedener Begabungsstufen andererseits bei bestimmten Tests. Für solche Vergleichungen lag schon früher einiges Material vor, auf welches Meumann seine Einteilung zum Teil stützt. Aber dies Material war meist bei Problemstellungen anderer Art mehr zufällig gewonnen und reicht daher für die Beantwortung unserer Frage noch nicht aus. Es müssen systematische Forschungen angestellt werden, die von vornherein das zu prüfende Menschenmaterial sowohl nach Altersstufen wie nach Begabungsgraden zweckmäßig auswählen, so daß der Leistungsvergleich nach beiden Dimensionen einwandfrei erfolgen kann.

Solche Untersuchungen sind entweder an abnormen oder an normalen Kindern anzustellen.

In die erste Gruppe gehören vor allem die gründlichen Untersuchungen von Chotzen (*53*) an Hilfsschulkindern; sie haben sehr deutlich die reinen Alterstests hervortreten lassen. Wie zu erwarten war, überwiegen unter diesen „Alterstests" solche, die sich auf häufig geübte Leistungen und Alltagserfahrungen beziehen, dagegen erweisen sich als eigentliche Begabungstests solche, die Neues darbieten, Ungewohntes verlangen und eine scharfe Aufmerksamkeitsspannung beanspruchen. Hierüber wird an anderer Stelle (S. 187) ausführlicher zu berichten sein.

Als Untersuchungen der zweiten Gruppe seien die von Gaßmann und Schmidt (*62*) und von Minkus und Stern (*31*) genannt.

Die Erstgenannten prüften die Merkfähigkeit mit Sätzen verschiedener Länge. Fassen wir die schon früher (S. 74/75) besprochenen Ergebnisse unter

unserem Gesichtspunkt zusammen, so ergibt sich: der Auffassungsumfang (Länge der noch richtig reproduzierbaren Sätze) variiert sehr stark mit dem Alter, aber wenig mit der I. Die Streuung dieses Umfangs (Unabhängigkeit von der silbenmäßigen Länge des Satzes) ist durchaus nicht eine Funktion des Alters, wohl aber eine solche der I.: die Begabtesten haben die größte Streuung. Der prozentuale Anteil der sinnvollen Fehler unter allen Fehlern endlich nimmt sowohl mit dem Alter wie mit der Begabung stark zu. Dies letzte Moment würde also bei einem Staffelsystem besonders zu berücksichtigen sein — mehr als der bisher vornehmlich benutzte Merkumfang, der ganz überwiegend Entwicklungs- und nicht Begabungstest ist.

Der von Minkus ausgearbeitete Test der Bindewortergänzung (siehe S. 30/31) besitzt, wie die von mir bearbeiteten Ergebnisse eines Massenversuchs zeigen, für die Jahrgänge 12—14 ebenfalls die doppelte Korrelation zum Alter und zu der Begabung, so daß auch er als „Begabungs-Alters-Test" angesprochen werden kann.

Mit der Methode der Scheidung zwischen Entwicklungs- und Begabungstests beschäftigt sich auch Claparède (*55*). Er bemängelt mit Recht, daß Meumann den Versuch macht, die Zugehörigkeit von Tests zur einen oder anderen Gruppe a priori zu bestimmen. Die Zuweisung sei vielmehr nur durch massenstatistische Berechnungen der Testleistungen verschiedener Altersstufen zu vollziehen.

Bei Tests mit abstufbaren Leistungen müsse man für jede Altersstufe die Streuungskurve feststellen; nur dann, wenn sich die Streuungskurven zweier aufeinanderfolgender Altersstufen ganz wenig kreuzen, haben wir das Recht, von charakteristischen Leistungen der einzelnen Jahrgänge, mithin von Entwicklungstests zu sprechen. Claparède schlägt als Maßstab vor, daß sich die Mittelwerte aufeinanderfolgender Jahresleistungen mindestens um das Vierfache des Streuungswertes jeder Kurve (gemessen an der „wahrscheinlichen Abweichung") unterscheiden müssen.

Die nicht abstufbaren Tests des BS.-Systems, bei denen es lediglich die Wertungen + oder — gibt, will Claparède nur dann als „Entwicklungstests" anerkennen, wenn in zwei aufeinanderfolgenden Altersstufen die Häufigkeit der richtigen Lösungen von 10% auf 80% steigt. Man sieht an dieser Angabe, daß auch Claparède von aprioristischen Erwägungen nicht ganz frei ist; in der Praxis der Testprüfungen kommen Altersfortschritte von solcher Sprunghaftigkeit nur in verschwindend seltenen Fällen vor.

6. Die Punktskalamethode.

Der Amerikaner Yerkes (*125*) hält sich zwar in der Auswahl der Tests eng an die Binetsche Liste, unterwirft aber das Berechnungsverfahren einer völligen Umgestaltung. Er stößt sich

zunächst an der groben Alternativwertung bei Binet, nach der für jeden Test nur ein „gelöst" oder „nicht gelöst" (+ oder —), aber keine Abstufung des Leistungsgrades möglich ist. Statt dessen führt Y. eine Punktwertung ein. Beispiele:

Bei dem Binetschen Test „60 Worte in 3 Minuten sagen" war es zweifellos ungerecht, daß 20 Worte ebenso als Nichtlösung bewertet wurden wie 58 Worte, 100 Worte ebenso nur als einfache Lösung wie 60 Worte. Yerkes beziffert dagegen: 30—44 Worte in 3 Minuten als einen Punkt, 45—59 als 2, 60—74 als 3, darüber als 4 Punkte. Oder: Bei den „absurden Sätzen" rechnet Binet als +, wenn 2 von dreien gelöst sind. Yerkes gibt 5 Absurditäten und wertet jede Einzellösung als einen Punkt.

Freilich wird hierdurch die einfache Berechnung des IA. unmöglich gemacht, und Yerkes muß einen anderen Maßwert an dessen Stelle setzen. Er richtet 20 Tests der Binet-Skala so ein, daß sie im Höchstfall, also bei bestmöglicher Lösung jedes Tests, 100 Punkte ergeben, und berechnet nun für jeden Prüfling die Prozentzahl der erlangten Punkte.

Damit verzichtet nun freilich Yerkes ganz auf die absoluten Maße, die wir vermittels des IA. im IQ. gewonnen haben. Er hat statt dessen die Absicht, mit seinem Hundert-Punkt-System für alle möglichen Kindergruppen verschiedenen Alters, Geschlechts, sozialen Standes usw. Normalwerte zu schaffen, an denen dann das individuelle Kind einer solchen Gruppe gemessen werden kann. Aber zur Herstellung von Normalwerten für ganz verschiedene Gruppen ist die Methode wegen ihrer Starrheit sehr wenig geeignet. Sämtliche Kinder nämlich, welchen Alters auch immer, werden danach beurteilt, wieviel sie von den insgesamt 100 Punkten des Systems erreicht haben. Die Folge ist, daß die kleinen Kinder überhaupt nur mit einer sehr geringen Anzahl von Tests geprüft werden, da die weitaus meisten Tests des Systems ihre Fähigkeit übersteigen. Somit ist das Bild der I., das man von den jüngeren Kindern erhält, ein viel unvollkommeneres als das der älteren, eigentlich ein wesentlich nur negatives; es wird überwiegend festgestellt, was sie nicht lösen können, nicht aber, was sie lösen können, da die ihnen adäquaten Aufgaben nur sehr spärlich sind. Daher besteht auch bei ihnen die Möglichkeit nur in geringem Maße, etwaige Ausfälle auf einem I.-Gebiet durch positive Leistungen auf einem anderen zu kompensieren. Hierin ist das BS.-System der Punktzahlmethode offenbar überlegen.

III. Teil.

Anwendungen und Ergebnisse. Intelligenz- und Schule.

VIII. Kapitel.

Gesamtverteilung der Intelligenzen.

Bereits vor einem Menschenalter hat Galton (*141*) die Vermutung ausgesprochen, daß die Begabungsgrade einer großen ungesiebten Menschenmasse nach der Gaußschen Kurve symmetrisch verteilt sein müßten. Die Gaußsche Kurve stellt das einfachste Verteilungsgesetz dar, nach welchem die um einen Mittelwert herumliegenden Fälle in größter Häufung auftreten, während die Abweichungen nach oben und nach unten um so seltener vorkommen, je weiter sie vom Mittelpunkt entfernt sind. Die Kurve hat die Form einer symmetrischen Glocke[1]). Durch den Wert der „wahrscheinlichen Abweichung" wird die Gaußsche Verteilungskurve in drei Fraktionen geteilt, derart, daß die Mittelfraktion (zwischen den wahrscheinlichen Abweichungen nach oben und unten) die Hälfte aller Fälle, die beiden Seitenfraktionen je 25% enthalten.

Galton — und im Anschluß an ihn Pearson (*160*) — hatten den Satz von der symmetrischen Normalverteilung der Intelligenzen als eine Art apriorischen Gesetzes aufgestellt und aus ihm eine Einteilung der vorkommenden Begabungen in eine Reihe von Stufen mit bestimmten Häufigkeitsprozenten abgeleitet. Pearson hatte dann diese Gliederung der Begabungsgrade durch eine nichtexperimentelle Erhebung zu stützen gesucht, indem er die I. von 4600 Schulkindern durch ihre Lehrer nach einer siebengradigen Skala schätzen ließ, ferner auch 1000 Studenten nach ihren üblichen Prüfungszeugnissen klassifizierte. Unter gewissen rechnerischen

[1]) Vgl. Differentielle Psychol. (*12*) S. 242ff.

Voraussetzungen fügte sich die Verteilung im wesentlichen der Gaußschen Kurve[1]).

Nunmehr liegt aber mit Hilfe der IP. eine empirische Bestätigung des Gesetzes vor, die so weit geht, als man unter den verwickelten Verhältnissen dieser Methodik nur erwarten darf. Der Nachweis dieser massenpsychologischen Gesetzmäßigkeit darf als das gewichtigste theoretische Ergebnis angesehen werden, das uns das Verfahren der Staffelmethode bisher geliefert hat; aber auch die praktischen Folgerungen, die daraus gezogen werden können, sind bedeutend genug.

Die experimentellen Belege für das Verteilungsgesetz gliedern sich in zwei Gruppen: eine Reihe älterer Untersuchungen, die sich im wesentlichen der Binetschen Testanordnung bedienen und durch deren Unvollkommenheit stark beeinträchtigt sind, und zwei neuere, welche mit einer verbesserten Testanordnung arbeiten und mit ihr zu ganz eindeutigen Ergebnissen kommen.

Die Ergebnisse der ersten Gruppe sind sämtlich in Tab. I

Tab. I.

Prüfer	Ort	Zahl der Prüflinge	Alter der Prüflinge	Abstand des IA. vom LA. in Jahresgraden				
				−2 und mehr	−1	0	+1	+2 und mehr
				%	%	%	%	%
Binet	Paris	203	?	6	21½	51	20½	1
Bobertag .	Breslau	261	5—10 J.	4	19	52	22½	2½
Goddard . .	Vineland	1277	5—11 J.	11	20½	41½	21½	5½
Wiersma .	Holland	141	6—12 J.	24		54	22	
Dougherty	Kansas City	290	6—11 J.	26		47	27	
Brigham . .	Princeton	162	6—11 J.	31		45	24	
	Insgesamt	2334	5—12 J.	27,2		48,4	24,4	

zusammengestellt. Die Untersucher bedienten sich der groben Differenzformel IA.—LA., wobei sowohl für LA. wie für IA. nur ganzzahlige Werte genommen wurden. Es wurde also lediglich festgestellt, wie viele Kinder auf dem Niveau ihres Alters standen („Niveaukinder") und wie viele einen I.-Rückstand bzw. -Vorsprung um 1, 2 . . . Jahre hatten. In der Tabelle sind die Kinder verschiedener Lebensalter vereinigt. Die Schlußreihe verbindet die Ergebnisse aller Untersucher zu Gesamtziffern.

[1]) Näheres über diese nichtexperimentellen Untersuchungen bei Jaederholm (72).

Man erkennt deutlich, daß die Verteilung überall eine wenigstens annähernde Symmetrie zeigt; auch die oben erwähnte Forderung der Gaußschen Kurve, daß die Mittelgruppe die Hälfte aller Fälle umfasse, ist mit Annäherung bestätigt. Im ganzen tritt ein geringes Überwiegen der unterwertigen gegenüber den überwertigen Kindern hervor, doch kommt diese Asymmetrie bei der Ungenauigkeit der angewandten Methode nicht in Betracht.

Es zeigt sich nämlich bei weiterer Gliederung der obigen Ergebnisse, daß die normale Verteilung nur gilt, solange die Kinder der verschiedenen Lebensalter zusammengeworfen werden, daß sie dagegen für die einzelnen Lebensalter nicht zutrifft. Dies erweist die folgende Tab. II für vier der obigen Untersucher:

Tab. II.

Prüfer	Alter in Jahren	Prozentzahl der Kinder unter dem Niveau ihres Alters		
		unter	auf	über
Bobertag	5—7	9½	57	33½
	8—9	13½	56	30½
	10—11	40½	40½	19
Wiersma	6—7	9	64	27
	8—9	23	50½	26½
	10—12	41	46	13
Goddard	5	12	35	53
	6	20½	30	49½
	7	13	58	29
	8	44	41	15
	9	40	28	32
	10	27½	56	16½
	11	56	36	8
Daugherty	6—7	1	55	44
	8—9	30	50	20
	10—11	41½	35½	23

Man sieht, daß bei den einzelnen Altersstufen von der symmetrischen Verteilung so gut wie nichts mehr zu bemerken ist. Nur bei der Altersstufe 8—9 schimmert sie noch zuweilen durch; dagegen zeigen die jüngeren Jahrgänge durchweg eine starke Verschiebung der Häufigkeit zu den über dem Niveau stehenden Kindern hin, die älteren ebenso durchgängig ein Überwiegen der unterwertigen Kinder. Die oben gefundene allgemeine Symmetrie ist also nur durch eine Ausgleichung der entgegengesetzten Asymmetrien in den Verteilungen der verschiedenen Jahrgänge zustande

gekommen; und man begreift nicht, wie Goddard diese Verhältnisse mißdeuten und auf Grund seiner Ziffern die symmetrische Begabungsverteilung bereits als erwiesen ansehen konnte.

Andererseits darf die bei den einzelnen Altersstufen zu findende Asymmetrie nicht als eine Widerlegung des Galtonschen Satzes angesehen werden; sie ist vielmehr lediglich eine Folge der falschen Testeichung: die von Binet für die niederen Altersstufen bestimmten Tests sind großenteils für die betreffenden Jahrgänge zu leicht, die für ältere Kinder sind großenteils zu schwer. Um diesen Grundfehler der Binetschen Anordnung zu beseitigen, bedarf es nicht eines völligen Ersatzes der zu leichten (bzw. zu schweren) Tests durch neue, sondern nur einer veränderten Eichung, d. h. einer Verschiebung der vorhandenen Tests im Staffelsystem. Wie schon S. 142 nachgewiesen, hat diejenige Testanordnung für eine Altersstufe als die angemessene zu gelten, für welche im Gesamtdurchschnitt aller Prüfungen das Intelligenzalter mit dem Lebensalter zusammenfällt; also die Gleichung IA.—LA. = 0 besteht. Erst auf der Grundlage dieser Gleichung kann dann die Frage aufgeworfen werden, wie sich sämtliche Kinder jeder Altersstufe um diesen Durchschnittswert herum verteilen.

Zwei neuere Forscher, Jaederholm (*72*) und Terman (*119*), haben nun die BS.schen Staffelberechnungen so lange abgeändert, bis sie für jede Altersstufe jene Grundgleichung IA.—LA. = 0 erreicht hatten. Sofort trat die Gültigkeit der Gaußschen Verteilungskurve mit Deutlichkeit hervor.

Jaederholm prüfte 200 Stockholmer Kinder, die teils 8, teils 10 Jahre alt waren. Er begnügte sich für LA. und IA. nicht mit ganzen Zahlen, sondern gewann genauere Werte derart, daß die Differenz IA.—LA. in sehr verschiedenen Abstufungen vorkam. Er schied nun nach gleichen Abständen dieser Differenz 20 Begabungsgruppen, auf welche sich die 200 Kinder in der in Tab. III angegebenen, „beobachteten" Häufigkeit verteilen. Daneben stehen die ideellen, „berechneten" Häufigkeiten, die sich bei absoluter Gültigkeit des Gaußschen Gesetzes ergeben würden.

Die Vergleichung beider Reihen zeigt eine weitgehende Übereinstimmung; der Grad dieser Übereinstimmung wird von J. nach einer Methode von Elderton (*140*) auf P. = 0,9 berechnet (wenn P. = 1 absolute Deckung bedeutet). Jaederholm zieht daraus den Schluß: „daß die Gaußsche Kurve mit großer Genauigkeit für die Verteilung der I.-Grade dieser 200 Kinder gültig ist".

Tab. III.

Begabungsgruppe	Beobachtete Häufigkeit	Berechnete Häufigkeit bei idealer Gauß-scher Verteilung
1	1	0,4
2	2	0,8
3	1	1,9
4	3	3,6
5	6	6,5
6	12	10,3
7	18	14,6
8	15	19,8
9	22	22,8
10	29	25,3
11	26	24,0
12	14	21,9
13	18	17,5
14	14	12,3
15	11	8,3
16	4	4,8
17	2	2,7
18	1	1,3
19	1	0,6
20	0	0,2
	Summa 200	

Der von Jaederholm ausgesprochene Wunsch, die Verteilung an einem weit größeren einwandfrei zusammengestellten Menschenmaterial nachzuprüfen, wird in der wichtigen Arbeit von Terman-Lyman-Ordhal (*171*) erfüllt. Es wurden 905 Kinder im Alter von 5—14 Jahren mit der sog. Stanford-Revision der BS.-Methode geprüft, bei der durch Verschiebung der Tests für jede Altersstufe die Gleichung IA.—LA. = 0 erfüllt war. Die Kinder waren nicht willkürlich ausgewählt; vielmehr wurden alle Kinder eines bestimmten Stadtteils, der eine durchschnittliche soziale Struktur hatte, herangezogen. Gründliche Vorübung der Prüfer und wörtliche Niederschrift der Antworten leisteten für eine sehr exakte Durchführung der Untersuchung Gewähr. Ein bedeutender Vorzug vor Jaederholm war die Benutzung des Intelligenzquotienten $IQ. = \frac{IA.}{LA.}$ an Stelle der Differenz IA.—LA. zur Feststellung der Verteilung. Denn diese Differenz ist, wie schon mehrfach betont, für verschiedene Lebensalter von verschiedener Bedeutung, gibt also nicht ein vom Lebensalter unabhängiges Bild der Verteilung[1]).

[1]) Da Jaederholm in seiner Untersuchung 8—10jährige Kinder vereinigt hatte, war die Benutzung jener ungleichwertigen Differenz erst recht nicht ohne Bedenken.

Dagegen zeigt die Termansche Arbeit so recht den Wert des IQ.-Maßes, da nun das Verteilungsbild für jedes Lebensalter ein durchaus analoges wird. T. bildet 9 Begabungsstufen, indem er immer 10 Hundertstel des IQ. zu einer Stufe vereinigt; die mittlere Stufe umfaßt die IQ. von 0,96—1,05. Tab. IV faßt die

Tab. IV.

IQ.	0,56 bis 0,65 %	0,66 bis 0,75 %	0,76 bis 0,85 %	0,86 bis 0,95 %	0,96 bis 1,05 %	1,06 bis 1,15 %	1,16 bis 1,25 %	1,26 bis 1,35 %	1,36 bis 1,45 %
117 6-jähr.	0	1	6	17	38	24	9	6	1
113 9 „	0	2	12	21	38	23,5	6	1	1
98 13 „	1	7	11	22,5	33	18,5	6	1	0
905 5-14 „	0,33	2,3	8,6	20,1	33,9	23,1	9	2,3	0,55

von Terman gegebenen Zahlen zusammen; sie stellen zuerst die Verteilung auf drei ziemlich weit auseinanderliegende Altersstufen, zuletzt die Gesamtverteilung der 905 Prüflinge dar. Die Annäherung an die Symmetrie und an die Gesetzmäßigkeit der Gaußschen Kurve ist in jeder Reihe sehr groß, insbesondere, wenn man die relativ kleine Zahl der Prüflinge jeder Altersstufe in Betracht zieht. Geradezu überraschend aber ist die Ähnlichkeit der drei verschiedenen Altersgrade untereinander; denn diese ermöglicht uns, nicht nur über die Form der Häufigkeitsverteilung, sondern auch über die absolute Streuungsbreite der Intelligenzen etwas Allgemeingültiges auszusagen.

Werden nämlich die 50% mittleren Intelligenzen vereinigt, so umfassen diese bei den verschiedenen Lebensaltern immer die ungefähr gleichen Breiten, nämlich die IQ. zwischen 0,94 und 1,07. Genauer lauten diese Ziffern[1]):

für die	6	jährigen	0,94	—1,07
„ „	8	„	0,94	—1,07
„ „	10	„	0,943	—1,067
„ „	12	„	0,938	—1,062
„ „	14	„	0,925	—1,075

Über die Verteilung der zu beiden Seiten der mittleren 50% liegenden je 25% der guten und schwachen Intelligenzen geben einige Ziffern bei Terman-Lyman, S. 557, Auskunft. Durch Vereinigung von verschiedenen Angaben gewinnen wir folgendes Schema der I.-Streuung:

[1]) Gewonnen durch Umrechnungen der Tabellen von *171*, S. 558.

Tab. V.

IQ.		umfassen	
unter 0,70	unter 0,77	1%	3%
0,70—0,76		2 „	
0,77—0,85	0,77—0,93	7%	22%
0,86—0,93		15 „	
0,94—1,00	0,94—1,07	25%	50%
1,01—1,07		25 „	
1,08—1,15	1,08—1,24	15%	22%
1,16—1,24		7 „	
1,25—1,30	über 1,24	2%	3%
über 1,30		1 „	

Hieraus können wir nun wieder für gleich große Gruppen der Kinder die dazugehörigen mittleren IQ. ableiten und auf solche Weise von dem einzelnen Kind auf Grund seines IQ. aussagen, welchen Rang es der I. nach innerhalb seiner Altersgenossen einnimmt. Gliedern wir die Gesamtheit der zu vergleichenden Kinder in 10 gleich umfangreiche Gruppen, so ergibt sich Tab. VI.

Tab. VI.

Erstes	Zehntel:	IQ. um	1,20
Zweites	„	„ „	1,12
Drittes	„	„ „	1,07
Viertes	„	„ „	1,03
Fünftes	„	„ „	1,01
Sechstes	„	„ „	0,99
Siebentes	„	„ „	0,97
Achtes	„	„ „	0,93
Neuntes	„	„ „	0,88
Zehntes	„	„ „	0,80

Da die Jaederholmschen Ergebnisse nach einem anderen Verfahren berechnet sind als die Termanschen, so läßt sich nicht sofort ersehen, wieweit sie auch in den absoluten Werten übereinstimmen. Ich habe deshalb aus den Rohtabellen von Jaederholm die Umrechnung in IQ. vorgenommen und diese entsprechend der Termanschen Tab. 4 geordnet (Tab. VII). Die Ziffern sind in

Tab. VII.

	IQ.	0,76 bis 0,85	0,86 bis 0,95	0,96 bis 1,05	1,06 bis 1,15	1,16 bis 1,25
100	8jähr.	4%	18%	52%	23%	3%
100	10jähr.		30%	48%	20%	2%
200	8 u. 10jähr.	2%	24%	50%	21,5%	2,5%

mancher Hinsicht lehrreich. Zunächst ist die Symmetrie, welche Jaederholm durch Vereinigen der beiden Altersstufen fand, in jeder einzelnen Altersstufe nicht so ideal verwirklicht — was ohne weiteres durch die geringe Zahl der Prüflinge verständlich wird. Immerhin ist auch so die Annäherung an die Symmetrie noch groß genug, um von einer Gültigkeit des Gaußschen Verteilungsgesetzes sprechen zu lassen. Sodann ist die Streuungsbreite der Differenzierung bei Jaederholm merklich geringer als bei Terman. Die IQ. zwischen 0,96 und 1,05 umfassen bei Jaederholm 50%, bei Terman nur 34% aller Fälle, die Fälle IQ. $<0{,}76$ und $>1{,}25$ fehlen bei Jaederholm ganz. Dies erklärt sich wenigstens für die schwächsten Intelligenzen einwandfrei daraus, daß Jaederholm nur normale Kinder unter Ausschaltung der Hilfsschulkinder, Terman aber völlig ungesiebtes Material geprüft hatte. Im übrigen aber kann die geringere Streuung entweder darauf beruhen, daß die schwedischen Kinder (schon wegen der viel größeren Einheitlichkeit der Rasse) intellektuell untereinander gleichartiger sind als die amerikanischen Kinder — oder daraus, daß die Jaederholmsche Abwandlung der BS.-Methode weniger deutlich differenziert als die Terman-Revision. Vermutlich werden beide Gründe beteiligt sein.

Immerhin ergibt sich doch aus dieser Vergleichung, daß die absoluten Streuungswerte, die wir für Terman berechneten und die wir zur Rangierung des einzelnen Kindes empfahlen, nicht ohne weiteres auf andere Verhältnisse übertragbar sind. Es ist eine der dringlichsten Aufgaben, daß wir in Deutschland — und zwar an verschiedenen Stellen — mit einer revidierten Staffelmethode und unter gleich sorgfältigen Vorsichtsmaßnahmen, wie es Terman tat, Massenprüfungen an mindestens je tausend Kindern vornehmen; erst dann werden wir das Recht haben, für jedes weitere zu praktischen Zwecken geprüfte Kind aus seinem IQ. auch seinen intellektuellen Rangplatz abzuleiten.

Ziehen wir nun aus dem gefundenen Verteilungsgesetz die theoretischen und praktischen Folgerungen.

1. In einer großen ungesiebten Menschenmasse verteilen sich die I.-Grade so, daß eine eingipflige, symmetrische Häufigkeitskurve entsteht. Das Maximum der Häufigkeit fällt also zusammen mit dem Durchschnittswert, der aus sämtlichen I.-Graden berechnet wird, und mit dem mittleren Rangplatz bei gradweiser Anordnung aller Individuen (dem sog. „Zentralwert").

2. Die Symmetrie der Verteilungskurve bedeutet, daß sich von dem normalen Mittelmaß aus die Individuen zu den Übermittel- und Untermittelwerten hin gleichmäßig verteilen. Es ist demnach nicht berechtigt zu behaupten, daß es mehr geistig schwache als geistig hochbefähigte Individuen gibt. Alle derartigen Behauptungen beruhen auf willkürlichen und ungleichmäßigen Abgrenzungen der beiden nichtnormalen Gruppen.

Dies gilt z. B. von einer Aufstellung Goddards, nach welcher 21% der Kinder unternormal, aber nur 4% übernormal sein sollen. Diese Ziffern sind leider unbesehen von Meumann (*9*) übernommen worden und seitdem in zahlreiche deutsche Darstellungen als feststehendes Ergebnis übergegangen. Und **dabei sind sie grundfalsch!**

Wie war nämlich Meumann zu jenen sonderbaren Häufigkeitsziffern gekommen? Ein Blick auf die von ihm zitierte Tabelle Goddards gibt Aufklärung (Bd. II, S. 198). Goddard hat hier die Prüfungsergebnisse von 2000 Kindern von 4 bis zu 15 Jahren zusammengefaßt, obgleich für die drei ältesten Jahrgänge im BS.-System gar keine Tests existieren, vermittels deren die Kinder ihr adäquates, geschweige denn ein höheres IA. erweisen könnten. So mußten unverhältnismäßig viel Kinder unter dem Niveau und verhältnismäßig wenig über dem Niveau stehen; die von Meumann übernommenen Ziffern beruhen also auf einem elementaren Rechenfehler!

3. Die Übergänge von der Durchschnittlichkeit zur Über- und Unternormalität sind durchaus stetig. Unter dem Gesichtspunkt des bloßen Gradunterschiedes gibt es im Verteilungssystem der Individuen keine Stelle, an der man eindeutig sagen könnte: hier hört die Normalität auf und beginnt die Unternormalität bzw. Übernormalität. Alle Gruppenbildungen, die man lediglich auf Grund gradueller I.-Feststellungen macht, sind künstlich.

4. Da aber solche künstlichen Gruppenbildungen innerhalb einer großen Masse von Individuen nicht zu entbehren sind, so ist folgendes zu beachten. Um den mittleren I.-Grad herum liegen die Fälle in großer Häufung und Ähnlichkeit dicht beieinander, wogegen nach den Seiten der Verteilungskurve die Häufigkeit immer spärlicher, die Verschiedenheit der Fälle immer deutlicher wird. Es ist nun natürlich zu vermeiden, die Schnitte in die Gegend größter Dichtigkeit und Ähnlichkeit zu legen; denn dann ist die Willkür besonders groß, durch welche ganz ähnliche Fälle verschiedenen Gruppen zugewiesen werden. Aus diesem Grunde ist zunächst die Halbierung der Gesamtmasse in eine „gute" und eine „schlechte" Hälfte besonders unglücklich. Aber auch die Drittelung in drei gleich zahlreiche Gruppen ist verfehlt; denn dann werden noch immer viele Fälle, die eigentlich in

das mittlere Hauptdichtigkeitsgebiet gehören, in die „gute" und „schlechte" Abteilung abgedrängt. Dagegen ist eine einigermaßen angemessene Abgrenzung des eigentlichen Mittelgutes durch die Abspaltung der beiden äußeren Viertel von der mittleren Hälfte zu erreichen und auch schon vielfach mit gutem Erfolg angewendet worden. Die Grenzstriche fallen dann mit den „wahrscheinlichen Abweichungen" der Gaußschen Kurve zusammen. Auf Grund der bisherigen Messungen würde das so definierte Mittelgut zwischen den IQ. 0,94 und 1,07 nach Terman (0,96 und 1,05 nach Jaederholm) liegen. Die beiden übrigen Gruppen von je 25% der Individuen können wir dann mit Recht als die „schwachen" und „starken" Intelligenzen bezeichnen. Dagegen wäre es verfehlt, in ihnen die Gruppen der „unternormalen" und „übernormalen" zu sehen; denn mit diesen Worten bezeichnen wir solche Menschen, die sich qualitativ im günstigen oder ungünstigen Sinne von denjenigen unterscheiden, deren geistige Fähigkeit im allgemeinen den durchschnittlichen Anforderungen des Lebens angepaßt ist[1]); über solche qualitative Verschiedenheiten aber gibt unsere Verteilungskurve, die sich lediglich auf die quantitative Abstufung der I.-Grade bezieht, als solche überhaupt keine Auskunft. Es bedarf besonderer empirischer Prüfung, um festzustellen, welche Grade der I.-Schwäche als Merkmale pathologischer Beschaffenheit (geistiger Minderwertigkeit) und welche höchsten Stufen der I. als Zeichen einer die Norm überragenden Gesamtbefähigung zu gelten haben. Bei diesen später besonders zu besprechenden Fällen handelt es sich jedenfalls nur um ganz wenige Prozente (etwa 2—3%); die je 25% der „schwachen" und „starken" Intelligenzen liegen daher noch zum weitaus größten Teil im Bereich des „Normalen". Diese Feststellung ist wichtig gerade gegenüber der von medizinischer Seite oft beliebten Erweiterung des Begriffs des Abnormen, und gegenüber der weit verbreiteten Ansicht, als ob das „Normale" eine eindeutige, punktuell festgelegte Beschaffenheit sei. Die Normalität hat eine beträchtliche Breite; es gibt innerhalb ihrer außer dem eigentlichen Mittelgut noch die „Schwachnormalen" und die „Starknormalen". Und deshalb sollte die Differenzierung der Menschen nach Fähigkeiten, wie wir sie jetzt in Schule und Berufsleben erstreben, nicht lediglich die abnormen Fälle oben und unten —

[1]) Über die teleologische Bedeutung des Normalitätsbegriffs vgl. Diff. Ps. (*12*) S. 156—157.

z. B. durch Hilfsschulen einerseits, Gymnasien für hervorragend Befähigte andererseits — abspalten, sondern auch für die Schwachnormalen und Starknormalen besondere Vorkehrungen treffen.

Wir gewinnen durch diese Betrachtung also schließlich eine Fünfgliederung der Begabungsgrade mit folgenden ungefähren Prozentsätzen:

Abnorm Schwachbefähigte . .	etwa	3%,
Schwachnormale	„	22%,
Normales Mittelgut	„	50%,
Starknormale	„	22%,
Abnorm Hochbefähigte . . .	„	3%.

Das Schema deckt sich ungefähr mit den Gliederungen, die auf Grund der experimentellen Befunde von Binet (*3*, *4*), Bobertag (*51*), Goddard (vgl. Tab. I), Terman (*119*) (vgl. Tab. 5) und Jaederholm (*71*, *72*) (vgl. Tab. VII) gewonnen worden sind.

Es ist aber bemerkenswert, daß auch die rein praktische Übung, soweit sie es mit Abstufungen von Leistungsfähigkeiten zu tun hat, instinktiv zu ganz ähnlichen Einteilungsprinzipien zu greifen sucht. Insbesondere hat ja die Schule in ihrem Prädikats- und Zeugniswesen fortwährend derartige Einteilungen vorzunehmen; und da entspricht schon die Stufenleiter der üblichen Prädikatnummern von 5—1 der obigen Gruppierung. Die „3", die die normale Mittelleistung bezeichnet, kommt stets in weit größerer Häufigkeit vor als irgend ein anderes Prädikat; für die Bewertung der schwachen Leistungen ist die „4" weit häufiger als die „5", für die Bewertung der guten die „2" weit häufiger als die „1". An einer Statistik über 2772 Zeugnisse hat Bobertag nachgewiesen, daß hier in der Tat ähnliche Häufigkeitsverhältnisse vorliegen: besser als genügend waren 25,7%, genügend 50,8%, unter genügend 23,5% der Zeugnisse (*51*, II, Tab. IV).

Bobertag hat auch noch eine andere Folgerung aus obigen Zahlenverhältnissen gezogen, die seitdem allseitig übernommen worden ist. Bei Alternativaufgaben (mit Lösung oder Nichtlösung ohne Abstufung) muß man fordern, daß die Altersstufe, für welche die Aufgabe bestimmt ist, die obige Verteilung zeige; es müssen also 25% der Aufgabe gegenüber „schwachbefähigt" sein, d. h. sie nicht zu lösen vermögen. Daher ist eine Lösungsprozentzahl von 75% der Eichungswert für solche Aufgaben; Tests, die in einer ungesiebten Masse bestimmten Alters mehr als 25% Versager ergeben, sind zu schwer, solche mit weniger Versagern zu

leicht für dieses Alter. Da die Aufgaben des Binetschen Staffelsystems durchweg Alternativtests sind, so ist hier die 75%-Ziffer seitdem den Eichungen zugrundegelegt worden.

In anderer Form macht Lipmann (*198*) Gebrauch von der obigen Gruppierung. Will man bei Massenstatistiken die Grenzen bestimmen, innerhalb deren ein Zahlenwert (z. B. Alter, Leistungsgrad usw.) variiert, so ist die Angabe der äußersten Grenzfälle bei der Spärlichkeit und relativen Zufälligkeit ihres Auftretens wenig empfehlenswert. Deshalb zieht Lipmann vor, in solchen Fällen die Grenzen der 50% mittleren Fälle anzugeben. So bezeichnet er z. B. die Altersstreuungen einer Schulklasse durch die Ziffern 9; 9 — **8; 10** — 8; 7 — wobei die mittlere Altersangabe den Zentralwert, die beiden anderen die Grenzen der mittleren 50% der Schüler bedeuten. Auch bei Vergleichung der Ergebnisse verschiedener Gesamtgruppen pflegt er die Mittelgruppen für sich und ebenso die je 25%igen Seitengruppen gesondert für sich zu vergleichen.

IX. Kapitel.

Intelligenzuntersuchungen nichtnormaler Kinder und Jugendlicher nach der BS.-Methode.

1. Der diagnostische Wert der IP.

Als Binet (*2, 48*) sein Testsystem erdachte, hatte er in erster Linie die Prüfung nichtnormaler Kinder im Auge, damit solche rechtzeitig erkannt und den Hilfsklassen und Schwachsinnigenanstalten überwiesen werden könnten. B. hat auch mit Simon an einer größeren Zahl solcher Kinder seine Methode erprobt, aber leider keine ausführlichere Darstellung dieser Untersuchung gegeben; wohl aber zieht er Folgerungen aus seinen Versuchen, welche die Beziehung des Schwachsinns zu seiner Methode auf überaus einfache Formeln bringen (*5*, S. 113). Die eine These betrifft den I.-Rückstand und lautet: „Ich meinerseits bin der Ansicht, daß jeder zweijährige I.-Rückstand auf ein Zurückgebliebensein schwerer Art schließen läßt." Die andere bezieht sich auf den I.-Stillstand und besagt, daß Imbezille nicht das IA. von sieben Jahren, Debile nicht das von neun Jahren überschreiten.

In den folgenden Jahren sind ungezählte geistig minderwertige Kinder in verschiedenen Kulturländern mit dem Staffelsystem ge-

prüft worden. In den meisten Fällen handelte es sich hierbei um praktisch-diagnostische Untersuchungen, über deren Ergebnisse nichts an die Öffentlichkeit kam. Daneben gibt es eine kleinere Reihe von Untersuchungen, welche einen Fortschritt in der wissenschaftlichen Erkenntnis und der Anwendungsweise der Methode brachten. Hierbei stellte sich mit der Zeit heraus, daß Binets einfache Formeln nicht ausreichten und daß die Methode als allein zureichendes Maß der geistigen Abnormität nicht benutzt werden könne. Dagegen erwies sich ihr großer Wert zur Ergänzung und Unterstützung anderer diagnostischer Hilfsmittel und zur massenstatistischen Ableitung allgemeinerer Ergebnisse über die I.-Verteilung und -Entwicklung bei abnormen Kindern.

In Amerika stellte Goddard (*143*) 1910 eine weit angelegte Untersuchung an, über die er aber leider nur unzureichende Angaben vorlegt. Er hat die jugendlichen und erwachsenen Insassen seiner Anstalt für Schwachsinnige in Vineland N. J., fast 400 Personen, nach BS. 1908 geprüft; er gibt aber nur an, wie oft die einzelnen Stufen des IA. vertreten waren, unterläßt jedoch die Bezugnahme auf das Lebensalter, so daß der Grad der Rückständigkeit in keiner Weise aus seinen Tabellen hervorgeht. Seine Zahlen lassen daher nur wenige, später zu erwähnende Schlüsse zu.

1912 und 1913 veröffentlichte Kuhlmann (*77*, *78*) seine an 1300 schwachsinnigen Kindern vorgenommenen I-Messungen hauptsächlich zu dem Zweck, um aus ihnen eine Neuordnung des Binetschen Staffelsystems abzuleiten.

Die ersten gründlichen Untersuchungen, die zugleich wirklich psychologische Erträge brachten, waren die Breslauer Prüfungen der Psychiater Kramer (*155*, *156*) und Chotzen (*53*). Sie bedienten sich der im Breslauer psychologischen Seminar von Bobertag ausgearbeiteten deutschen Form der Staffelmethode. Und zwar ergänzen sich die beiden Untersuchungen in erfreulicher Weise dadurch, daß verschiedenartige Kategorien von Kindern geprüft wurden. Kramers Material bestand nämlich zum Teil aus Jugendlichen, die vor das Jugendgericht gekommen und von diesem der psychiatrischen Begutachtung überwiesen worden waren, zum anderen Teil aus solchen Kindern, welche wegen psychischer oder wegen nervöser Erkrankung die Klinik oder Poliklinik aufsuchten. Chotzen wandte (gemeinsam mit Nicolauer) die Methode als städtischer Hilfsschularzt an; er prüfte damit alle Kinder, welche neu der Hilfsschule überwiesen waren. Während es Kramer mehr mit höheren Jahrgängen zu tun hatte, brachte es die Natur der Aufgabe Chotzens mit sich, daß seine meisten Prüflinge im Alter von acht und neun Jahren standen, doch er-

weiterte er seine Untersuchung durch Hinzunahme auch älterer Hilfsschüler.

Es folgen dann eine Reihe weiterer Untersuchungen an Hilfsschulkindern, und zwar deutscherseits in Kattowitz durch Bloch und Lippa (*127, 128*) und in München durch Weigl (*174*), ferner in Genf durch Alice Descoeudres (*58, 137*), in Stockholm durch Jaederholm (*71*).

Die fortschreitende Forschung hat gezeigt, daß das Verhältnis der geistigen Minderwertigkeit zum Ergebnis der IP. durchaus nicht so einfach ist, wie es Binet noch glaubte. Zunächst war schon sein Rechnungsgrundsatz unzulänglich. Es ergab sich nämlich, daß das von ihm aufgestellte Maß des I.-Rückstandes (IR. = IA.—LA.) keine einheitliche Bedeutung habe; denn die von ihm für die Zuweisung zur Hilfsschule festgesetzten „zwei Jahre Rückständigkeit" haben bei verschiedenen Altersstufen ja einen ganz verschiedenen Sinn, sind also kein konstantes I.-Maß. Ein Rückstand von zwei Jahren ist beim sechsjährigen Kind ein Zeichen starker geistiger Minderwertigkeit; er verrät bei einem neunjährigen einen sehr viel geringeren Schwachsinnsgrad; und für ein zwölfjähriges Kind braucht die durch Zweijahrs-Rückstand ausgedrückte Schwäche überhaupt noch nicht jenseits der Grenze der Normalität zu liegen. Hierauf ist schon allgemein an früherer Stelle hingewiesen worden (S. 143); doch sei hier noch ausdrücklich die besondere Erfahrung eines Psychiaters angeführt. Kramer (*155*) sagt:

„Bei den uns von der Jugendfürsorgezentrale überwiesenen Kindern . . . zeigte es sich deutlich, daß bei den jüngeren Kindern meist nur geringe, in den höheren Altersstufen immer größere Differenzen sich ergaben, während die Defektzustände, soweit wir aus den sonstigen Kriterien ein Urteil darüber haben, durchaus keinen entsprechenden Unterschied zeigten, sondern im Durchschnitt etwa gleich zu sein schienen." Ähnliches stellt Chotzen fest (*53*, S. 493).

Dieser rechnerische Mangel der Messung läßt sich nun völlig beseitigen, indem man an die Stelle der Differenz IA.—LA. den Intelligenzquotienten $\frac{\text{IA.}}{\text{LA.}}$ setzt, der, wie schon früher (S. 143) erwähnt, das erforderliche Maß von Konstanz besitzt, um Messungen und Vergleichungen unter Absehung vom Lebensalter zu ermöglichen. Einige neuere Forscher benutzen schon dieses Maß; andere bleiben bedauerlicherweise noch bei der Benutzung der Differenz bestehen und verschlimmern den Fehler noch dadurch, daß sie die gleichnamigen Differenzen bei verschiedenen Lebensaltern zu

Durchschnittszählungen zusammenwerfen. Ich habe mich deshalb der Mühe unterzogen, überall, wo das veröffentlichte Material es zuließ, die Angaben über IA. und IR. in IQ. umzurechnen. So sind die Tabellen dieses Kapitels zustande gekommen.

Wie steht es nun aber mit dem sachlichen Gehalt von Binets Meinung, daß die IP. die geistige Abnormität zum eindeutigen und einwandfreien Ausdruck bringe? Hier ist gegenüber den sanguinischen Hoffnungen Binets und mancher Amerikaner eine gewisse Zurückhaltung vonnöten. Die Rückständigkeit der I. ist stets nur eine Teilerscheinung innerhalb der geistigen Anormalität; der Unterschied des abnormen vom normalen Kinde ist fast niemals allein in einer graduellen Minderung der geistigen Fähigkeit, sondern zugleich in gewissen qualitativen Verschiebungen und Disharmonien zu finden. Das Kennzeichnende des Hilfsschulkindes oder des Fürsorgezöglings ist die soziale Unzulänglichkeit; diese ist das Ergebnis sehr komplexer psychischer Bedingungen, unter denen die geistige Rückständigkeit nur eine an Bedeutung sehr wechselnde Teilbedingung zu sein braucht.

Diese Betrachtung wird uns davon zurückhalten müssen, lediglich auf Grund einer IP. eine endgültige Entscheidung über das Vorhandensein einer Abnormität und über die Zuweisung zu einer Hilfsschule (bzw. über eine sonstige Bestimmung der Lebensbahn) zu treffen. Hier werden die qualitativen Feststellungen, wie sie in den Urteilen des Lehrers über die Schulfähigkeit und in den ärztlichen Diagnosen vorliegen, stets ihre notwendige Bedeutung behalten. Aber neben diesen Methoden und in enger Wechselwirkung mit ihnen wird die IP. ein um so wertvolleres Untersuchungsverfahren sein können. Sie kann in ganz kurzer Zeit bei noch nicht bekannten Kindern ein vorläufiges Bild der I. geben; sie läßt oft die psychologischen Gründe erkennen für Defekte, die dem Lehrer oder Arzt nur als komplexe Erscheinungen bekannt werden; sie macht auf bisher noch nicht bemerkte Schwächen aufmerksam; sie erlaubt exakte Vergleichungen der Kinder untereinander und die rangweise Einordnung eines einzelnen Kindes in die Reihe seiner Genossen. Darum ist das Urteil Kramers über den Wert der Methode durchaus zutreffend:

„Als Gesamtergebnis möchte ich zunächst hervorheben, daß wir mit der Methode für unsere Zwecke recht zufrieden waren. Ganz abgesehen von den zahlenmäßigen Resultaten lernt man im Laufe der Prüfung infolge ihrer Vielseitigkeit das Kind in seiner Eigenart, in seinem Verhalten auf Anforderungen verschiedener Art in ausgezeichneter Weise kennen, und erhält für die qualitativ

abweichende Reaktionsweise des Schwachsinnigen gute Anhaltspunkte. Bei den von der Zentrale für Jugendfürsorge gesandten Kindern, die meist nur zu einer einmaligen Untersuchung zu bekommen waren, machte sich die relativ kurze Zeitdauer der Prüfung (etwa $^3/_4$ bis 1 Stunde), in der man zu einem verwertbaren Urteil über die Intelligenz gelangt, äußerst angenehm geltend. In allen den Fällen, in denen auf Grund anamnestischer Daten oder auf Grund klinischer Beobachtungen sich sonst ein Urteil über die I. gewinnen ließ, haben sich Widersprüche zu dem Ergebnisse der Binetprüfung mit ganz geringen Ausnahmen nicht ergeben." (*155*, S. 27.)

Neuerdings äußern sich auch Hilfsschullehrer mit Wärme über den Wert, den die BS.-Methode für ihre Zwecke besitzt. So deckt Kannegießer (*151*a) die großen Mißstände auf, die mit dem heute üblichen, ganz ungeregelten und ungleichmäßigen Aufnahmeverfahren verknüpft sind und betont die unbedingte Notwendigkeit eines exakten und gleichmäßigen Aufnahmeverfahrens für die Hilfsschule; als solches könne die BS.-Methode trotz ihrer Mängel dienen. — Schmidt (*163*a) gibt den Hilfsschullehrern Anweisungen, wie sie aus der Berücksichtigung der einzelnen Testleistungen und -ausfälle eine feinere psychologische Analyse ihrer Zöglinge gewinnen könnten.

Der oben ausgeführte Gedanke, daß der Grad der intellektuellen Rückständigkeit kein eindeutiges Kennzeichen sozialer Unzulänglichkeit sei, ist vor allem von Pearson und Jaederholm (*161*) betont worden. Sie wenden sich gegen die Meinung, als ob geistige Schwäche überhaupt eine einheitliche Eigenschaft (eine „Mendelsche Einheit") sei, durch deren Besitz oder Nichtbesitz eine scharfe Grenze zwischen Abnormen und Normalen gezogen werde; geistiger Defekt sei vielmehr eine in stetigen Abstufungen vorkommende Variante, wobei gleiche Grade oft genug bei Normalen und Nichtnormalen vorkommen können. Sie stützen sich hierfür auf vergleichende Untersuchungen, die Jaederholm an fast sämtlichen (301) Hilfsschulkindern Stockholms einerseits, an 261 normalen Volksschulkindern verschiedenen Alters andererseits vorgenommen hat. Die von mir auf IQ. umgerechneten Ergebnisse sind in Tab. VIII zusammengestellt, welche ergibt, wieviel Kinder absolut und prozentual auf die verschiedenen IQ. kommen.

Die Tabelle bestätigt die Behauptung von Pearson und Jaederholm insofern, als ein breites Gebiet von IQ. beiden Kindergruppen gemeinsam ist. Aber das Bild ändert sich, wenn man auf die Häufigkeitsverteilung achtet. Rund 90% aller Hilfsschüler haben IQ. zwischen 0,61 und 0,90, der Häufigkeitsgipfel liegt um

Tabelle VIII.

Verteilung von normalen und Hilfsschulkindern auf die verschiedenen IQ.

	IQ.	unter 0,50	0,51 bis 0,60	0,61 bis 0,70	0,71 bis 0,80	0,81 bis 0,90	0,91 bis 100	101 bis 110	111 bis 120	über 120	Summa
in absol. Zahlen	norm.	—	—	3	11	43	74	80	47	3	261
	Hilfsschüler	5	16	69	130	71	9	1	—	—	301
in %	norm.	—	—	1	4	16½	28½	31	18	1	100
	Hilfsschüler	1⅔	5½	23	43	23½	3	⅓	—	—	100

0,75 herum, d. h. bei der „Dreiviertels-Intelligenz"; dagegen gruppieren sich die normalen Kinder um den Gipfelpunkt IQ. = 1, und zwar sind 60% zwischen 0,91 und 1,10 gelegen, also außerhalb des Hauptgebietes der Abnormen. Gibt also auch die IP. kein eindeutiges Symptom für die Zuweisung an die Hilfsschule, so ist sie doch in hohem Maße geeignet, die anderweitige Diagnose, die auf ärztliche und schulische Minderwertigkeit lautet, zu stützen und zu ergänzen bzw. zu zeigen, daß jene Unzulänglichkeit auf anderen Gründen als auf geistigen Defekten beruhen muß. Auch dies kann unter Umständen eine wertvolle Einsicht bedeuten.

Einen besonders heftigen Gegner hat die psychopathologische Anwendung der Binet-Methode neuerdings in Ziehen (*179*, S. 188) gefunden. Er sagt. „Die Untersuchungsmethode von BS., die leider wegen ihrer schematischen Bequemlichkeit auch in Deutschland Eingang gefunden, ist nicht nur wissenschaftlich unbrauchbar und oberflächlich, sondern gibt auch praktisch direkt zu Fehldiagnosen Anlaß."

Nun wäre ja die Gefahr solcher Fehldiagnosen nur dann vorhanden, wenn man sich allein auf die BS.-Methode verließe — was von allen besonnenen Anwendern heute als unzulässig angesehen wird. Wie wertvoll die Methode aber gerade dann werden kann, wenn sie mit der ärztlichen Diagnose organisch verbunden wird, haben schon die Untersuchungen von Chotzen und Kramer erwiesen, die neuerdings besondere, für Psychiater bestimmte, Darstellungen gefunden haben (*134*, *156*); es ist aber doch sehr zu begrüßen, daß die Behauptung Ziehens jüngst von ärztlicher Seite einer besonderen Nachprüfung unterworfen worden ist — mit dem

Ergebnis, daß seine Bedenken im wesentlichen gegenstandslos sind. W. Sauerbrey (*163*) hat in der Heil- und Pflegeanstalt Klein-Meusdorf bei Leipzig 142 meist mit I.- und anderen Defekten behaftete Knaben und Mädchen im Alter von 9—13 Jahren geprüft und das Ergebnis mit den auf genauer persönlicher Kenntnis beruhenden ärztlichen und pädagogischen Beurteilungen verglichen. Er fand im großen und ganzen zwischen beiden Bewertungen Übereinstimmung. In 6 Fällen schien die BS.-Methode nicht richtig angezeigt zu haben; von diesen bekamen 4 nach der Prüfung ein zu hohes, 2 ein zu niedriges IA. Wenn man bedenkt, daß die Eichung des Staffelsystems für deutsche Verhältnisse noch keine endgültige ist, so muß man von vornherein mit Unstimmigkeiten rechnen und wird angenehm überrascht sein, daß diese nur in so geringer Zahl auftreten. Sauerbrey hält daher auch die gefundene Übereinstimmung für gut, das Verfahren für brauchbar. Selbst die beiden Fälle der Unterschätzung der I. hätten nicht zu schweren Fehldiagnosen führen können, da die Differenzzahlen zu niedrig waren, um zur Annahme pathologischen Schwachsinns führen zu können. Als ein Glied in der Kette der psychiatrischen Untersuchungsverfahren will daher S. die Methode angewandt wissen, schon deshalb, weil es erwünscht ist, daß es eine einheitliche und damit vergleichbare Prüfungsmethode gibt.

Die diagnostische Bedeutung der IP. wird nun aber noch von einer anderen Seite her beträchtlich gestützt, nämlich durch die wiederholte Prüfung identischer Kinder. Denn die Wiederholung scheint in den weitaus meisten Fällen die erste Diagnose, zum mindesten die der ersten Nachprüfung, zu bestätigen. Mit Hilfe des IQ. sind wir jetzt imstande, die Prüfungsergebnisse aus verschiedenen Altersstufen desselben Kindes auf ein vom Alter unabhängiges Maß zu bringen, und es zeigt sich, daß wir damit wirklich ein Dauermaß, also einen Ausdruck für die konstitutionelle Beschaffenheit seiner I.-Anlage gewonnen haben.

Solche wiederholten Prüfungen sind von Alice Descoeudres (*58*) an 24 Kindern der Genfer Hilfsklasse angestellt und von mir für die vorliegenden Zwecke umgerechnet worden. Die Kinder wurden — immer ungefähr in Jahresabständen — zwei- oder dreimal, je eins auch vier- und fünfmal geprüft. Ich glaube aus den Prüfungen folgendes Ergebnis ableiten zu können:

Neben einigen atypischen Fällen bestehen unter den Schülern der Genfer Hilfsklassen zwei Typen. Der eine Typ zeigt von

vornherein eine ziemlich große Konstanz des IQ.; hierbei darf praktisch von „Konstanz" gesprochen werden, wenn sich die Werte der verschiedenen Jahre in einem Spielraum bis zu 5% bewegen. — Der andere Typ gewinnt die Konstanz erst nach einem Aufstieg von der ersten zur zweiten Prüfung; da die erste Prüfung gewöhnlich beim Eintritt in die Hilfsklasse stattfand, so bedeutet jener Aufstieg die Empfänglichkeit für die pädagogischen Einflüsse der Hilfsklasse. Es handelt sich hier also offenbar um solche Kinder, deren an sich schwache I. durch frühere unzulängliche oder ungeeignete Schulung noch in stärkerer Latenz gehalten worden war; sie wird nun durch die dem Kinde angemessene Hilfsschule geweckt und steigt auf den ihr möglichen Bruchteil der Normal-I., der nun für längere Zeit gewahrt wird. Übrigens findet sich dieser schroffe Anfangsanstieg nur bei jungen Kindern, nämlich bei solchen, deren erste Prüfung in den Altersstufen 7;1—8;6 lag.

Ein Abfall des IQ. von größerem Umfang fand sich nur ganz selten und dann fast ausschließlich bei den höheren Altersstufen, bei denen schon aus methodischen Gründen damit zu rechnen ist.

Zur Beleuchtung dieser Ergebnisse mögen einige bezeichnende Werte dienen, die aus der von mir aufgestellten Tabelle nach Descoeudres ausgezogen sind (Tab. IX).

Tabelle IX.

Verlauf des IQ. bei einigen mehrfach geprüften Kindern.

Kind	geprüft im Alter	IQ.	Änderungen des IQ.
a	6;10	0,82	
	7;10	0,82	0
b	7;7	0,74	
	bis	0,87	+ 0,13
	9;7	0,86	— 0,01
c	9;3	0,50	
	bis	0,54	+ 0,04
	12;3	0,53	— 0,01
d	10;3	0,59	
	bis	0,64	+ 0,05
	13;3	0,64	0
		0,63	— 0,01
e	9;0	0,71	
	bis	0,75	+ 0,04
	13;1	0,69	— 0,06
		0,70	+ 0,01
		0,63	— 0,07

Der Fall b gehört dem Typ mit anfänglichem Aufstieg des IQ., die anderen mehr oder minder dem Konstanztyp an. Für die Fälle c, d, e darf man wohl sagen, daß hier bereits über eine 3—4jährige Zeitspanne hin ein einigermaßen stabiles I.-Maß festgestellt sei von rund $^5/_{10}$ bzw. $^6/_{10}$ bzw. $^7/_{10}$ der normalen Intelligenz.

Alice Descœudres (*137*) hat auch noch mit Hilfe einer anderen Methode, nämlich der Rangkorrelation (s. S. 65), den diagnostischen Wert von Testserien für schwachsinnige Kinder nachgewiesen. Sie berechnete nämlich die Übereinstimmung zwischen der geschätzten I. und der Abstufung der Testleistungen. Die untersuchten 14 Kinder einer Anstalt gehörten zwar sehr verschiedenen Altersstufen an (von $6^1/_2$ bis zu 14 Jahren), aber es war doch möglich, ihre I. nach dem Gesamteindruck, den sie zu Haus und im Unterricht machten, zu schätzen, und sie hiernach zu rangieren. Tabelle X enthält nun in der ersten Zahlenrubrik die Korr. der Einzeltests zu der geschätzten Rangordnung nach der Höhe geordnet; sie schwanken zwischen 0,88 und 0,51. Die Verfasserin berechnete aber auch die resultierende Rangordnung für die gesamte Testserie und fand hier die Korr. 0,99 zur Intelligenzschätzung, also fast völlige Übereinstimmung beider Reihen. Ob diese Zahl ganz korrekt berechnet ist, habe ich nicht festgestellt. Dagegen habe ich selbst nach den Rohtabellen die Rangplätze für je

Tabelle X. A. Descœudres: Versuche an schwachsinnigen Kindern.

Tests	Korrelationen der einzelnen Tests mit der ISch.	Korrelationen von je 5 vereinigt. Tests zur ISch.	Korrelationen von sämtl. vereinigt. Tests zur ISch.
*1. Vergleichung von Begriffen	0,878		
2. Rechnen	0,868		
*3. Bildbeschreibung	0,842		
*4. Intelligenzfragen	0,817	**0,91**	
5. Unterscheidung von Körnern mittels des Tastsinns	0,812		
*6. Definitionen	0,801		
7. Perlenaufziehen	0,780		
8. Einbildungskraft (ein Bild wird gezeigt: was sprechen die abgebildeten Leute?)	0,761	**0,84**	
9. Geduldspiel (zerschnittenes Bild zusammenzusetzen)	0,734		**0,99**
*10. Kenntnis von 4 Münzen	0,699		
11. Aufmerksamkeit (alle a in einem Text anstreichen)	0,671		
12. Visuelles Gedächtnis (5 eben gezeigte Objekte nennen)	0,646		
13. Erkennen lückenhafter Zeichnungen	0,637	**0,73**	
14. Auditives Gedächtnis (5 Worte wiederholen)	0,539		
*15. 60 Worte in 3 Min. aufzählen	0,509		

5 Tests kombiniert und stets ergaben sich starke Korrelationen zur geschätzten I., die beträchtlich höher waren, als die höchste Korr. der daran beteiligten Einzeltests (zweite Zahlenrubrik).

Man darf nun aber nicht glauben, daß sich in diesen individual-diagnostischen Leistungen die Bedeutung der IP. für die Kindes-Psychopathologie erschöpft; vielmehr bietet sie uns auf dem Wege der Massenstatistik eine Reihe von Einsichten allgemeinerer Art. Denn da sie eine überall gleichmäßig anwendbare Methode darstellt, ist eine weitgehende gruppenmäßige Vergleichung der Kinder und eine Erarbeitung von Durchschnittsergebnissen ermöglicht.

2. Massenuntersuchungen an Hilfsschulen.

Über den Vergleich von Hilfsschülern und Normalschülern haben wir schon S. 174 im Anschluß an Pearson und Jaederholm gesprochen. Jetzt seien andere Ergebnisse erwähnt, die innerhalb der Hilfsschule erzielt wurden.

Da fällt zunächst in die Augen, daß die durchschnittliche I. der Hilfsschulinsassen sich durchweg innerhalb eines engen Bereichs, zwischen den IQ. 0,67 und 0,77 bewegt. Tab. XI, in der die mir zugänglichen Hilfsschulprüfungen unter gleichmäßiger Umrechnung auf IQ. vereinigt sind, zeigt, daß diese Gesetzmäßigkeit international gilt; denn neben drei deutschen Hilfsschulen finden wir je eine aus der französischen Schweiz und aus Schweden. Hilfsschulkinder haben also durchschnittlich Zweidrittel- bis Dreiviertel-Intelligenz — welcher Satz natürlich nicht ausschließt, daß sich eine Reihe von Kindern von beträchtlich höherem und erheblich tieferem I.-Stande unter ihnen befinden. Solche bedeutenden Abweichungen vom Durchschnittswert finden sich auf allen Altersstufen.

Tabelle XI. Durchschnittliche IQ. von Hilfsschulkindern.

	Lebensalter								
	7	8	9	10	11	12	13	14	15
Chotzen (Breslau) . . .	—	0,79	0,72	0,70	0,67		—	—	—
Weigl (München) . . .	—	0,75	0,76	0,73	0,71	0,67	—	—	—
Bloch-Lippa (Kattowitz) .	—	0,74	0,72	0,69	0,68	0,68	0,62		—
Descœudres (Genf) . .	0,78	0,75	0,75	0,67	0,67	0,60	—	—	—
Jaederholm (Stockholm) .	—	—	0,71	0,77	0,77	0,77	0,73	0,74	0,71

Vergleicht man die Altersstufen, so ist mit den Jahren eine geringe Abnahme der durchschnittlichen IQ. festzustellen. Diese Erscheinung ist vermutlich auf zwei Gründe zurückzuführen, deren erster ganz äußerlicher Natur ist. In dem System von Binet und Bobertag sind bekanntlich die Tests noch nicht genau den Altersstufen angepaßt, vielmehr die für jüngere Jahrgänge bestimmten zu leicht, die der höheren zu schwer. Infolgedessen erhalten die älteren Kinder in der Berechnung nicht ganz das ihnen zukommende IA., und der IQ. wird zu tief. Daß dieser methodische Mangel stark mitspielt, geht aus den abweichenden Ergebnissen von Jaederholm hervor. Dieser hatte nach einer korrigierten Staffelreihe mit genau geeichten Tests gearbeitet; infolgedessen ist in den Jahrgängen von 10—12 keine Spur des Absinkens bemerkbar; der durchschnittliche IQ. bleibt völlig konstant; erst bei den ganz hohen Jahrgängen, für welche die vorhandenen Tests vielleicht nicht mehr voll ausreichten, ist wieder eine geringe Abnahme des IQ. vorhanden.

Bei den deutschen Hilfsschulkindern, vielleicht auch in Genf, mag dann noch ein schultechnischer Grund hinzukommen, der bewirkt, daß die älteren Jahrgänge der Hilfsschule im Durchschnitt stärkere Schwachsinnsgrade enthalten als die jüngeren[1]). Die Überweisung aus der Normal- zur Hilfsschule wird nämlich bei jüngeren Schülern leichter, bei älteren nur dann geschehen, wenn sie sich als hochgradig unfähig erweisen; auch sind bei kleineren, erst kurze Zeit in der Schule befindlichen Kindern viel eher Irrtümer über die Befähigung möglich. Umgekehrt werden jüngere Hilfsschüler, deren geistige Schwäche sich als nicht sehr stark herausstellt, später gelegentlich zur Normalschule zurück überwiesen, so daß in den höheren Hilfsschulklassen nur schwerere Schwachsinnsgrade übrigbleiben.

Bobertag hatte seinerzeit — als ihm lediglich die Zahlen von Chotzen vorlagen — die Abnahme des IQ. mit den Jahren anders gedeutet; er sah sie als Zeichen dafür an, daß der IQ. überhaupt kein konstanter Ausdruck für den I.-Grad sei, sondern unter gleichen I-Verhältnissen mit den Jahren langsam abnehme. Die neueren Untersuchungen von Descœudres (*58*), Jaederholm (*161*), Weigl (*174*) zeigen, daß diese Annahme irrig ist; denn bei wiederholter Prüfung identischer Kinder oder beim Vergleich von Altersgruppen, die intellektuell wirklich

[1]) Dies bestätigt sich auch in den ärztlichen Diagnosen Chotzens; denn die milden Beurteilungen („nicht-schwachsinnig", „debil-fraglich") traten bei 39% der 8jährigen auf, dagegen nur noch bei 20% und 15% der 9- und 10jährigen Kinder, die zur Aufnahme in die Hilfsschule gemeldet wurden.

gleichwertig sind[1]), findet sich nichts von sinkender Tendenz mit steigenden Jahren, sondern annähernde Konstanz (und sogar gelegentliche Neigung zu geringem Anstieg). Es ist deshalb nicht zulässig, aus obigem Befund ein Argument gegen die Verwendbarkeit des Quotientenmaßes abzuleiten.

Noch speziellere Gruppierungen der Hilfsschulkinder werden möglich auf Grund der Untersuchungen Weigls (*174*) an 93 Kindern einer Münchner fünfklassigen Hilfsschule. Die von mir zusammengestellte Tab. XII erlaubt nämlich, die IQ. nach drei Richtungen zu vergleichen.

Tabelle XII. Verteilung des IQ. auf Altersstufen und Klassen nach Weigl.
(Kl. I ist die unterste.)

Altersstufe	Klasse I		Klasse II		Klasse III		Klasse IV	
Jahr	Anzahl	IQ.	Anzahl	IQ.	Anzahl	IQ.	Anzahl	IQ.
8	8	0.74	7	0,77	(1)	(0,75)	—	—
9	7	0,70	5	0,73	12	0,80	—	—
10	(2)	(0,55)	8	0,69	6	0,72	9	0,80
11		(3) (0,55)			5	0,73	7	0,76
12	—	—	—	—	—	—	7	0,67
13	—	—	—	—	(1)	(0,54)	(2)	(0,58)

Die senkrechten Kolumnen enthalten je eine Klasse nach Altersstufen gegliedert. Wieder zeigt sich das Absinken des IQ. mit steigendem Alter, aber jetzt in viel höherem Maße, was ja auch ganz natürlich ist; denn wenn ein älteres Kind noch in derselben Klasse weilt, in der auch jüngere Kinder sind, so wird es an angeborener Begabung weit hinter jenen zurückstehen.

Die wagerechten Reihen zeigen, wie sich die gleichaltrigen Kinder auf die verschiedenen Klassen verteilen; hier nimmt der IQ. mit steigender Klasse zu (durchschnittlich von Klasse zu Klasse um ungefähr 5%). Auch dies Ergebnis entspricht pädagogischer Beobachtung, die hier aber zum ersten Male ziffernmäßig festgelegt wird.

Endlich können wir noch in den verschiedenen Klassen die „gleichwertigen" Gruppen vergleichen, nämlich die Kinder mit

[1]) Vgl. die sogleich zu besprechenden Ergebnisse Weigls.

gleichem Abstand zwischen Lebensalter und Klassengrad. Den 8jährigen in Klasse I entsprechen die 9jährigen in Klasse II und die 10jährigen in Klasse III; denn jene 8jährigen Anfänger werden ja voraussichtlich selbst mit 9 Jahren in Klasse II und mit 10 Jahren in Klasse III sein. Solche gleichwertigen Gruppen sind in der Tabelle durch die schrägen Linien verbunden; ihre IQ. zeigen keine eindeutige Zunahme oder Abnahme, sondern nur ganz geringe Schwankungen innerhalb eines Spielraums von etwa 5%, also praktisch Konstanz: die in den verschiedenen Klassen und Altersstufen vorhandenen Vertreter gleichwertiger Gruppen haben ungefähr gleiche durchschnittliche IQ. — hierin sehe ich die wichtigste Probe auf den Wert dieses Maßprinzips.

Insbesondere treten bei Weigl drei solcher „gleichwertigen" Gruppen hervor: diejenigen Kinder, welche erst mit 9 Jahren in die unterste Hilfsschulklasse kommen (IQ. um 0,70 herum); jene, die von 8 Jahren an die Hilfsschule mit Erfolg absolvieren (IQ. um 0,74 herum); diejenigen, welche mit 8 Jahren schon für den zweiten Jahrgang der Hilfsschule reif sind (IQ. um 0,78 herum).

Die gewonnenen Werte der drei Gruppen haben eine große Ähnlichkeit mit den durchschnittlichen IQ., die sich bei medizinisch zusammengehörigen Gruppen finden, worüber im nächsten Abschnitt berichtet wird.

3. Die verschiedenen Schwachsinnsgruppen.

Weitere massenstatistische Untersuchungen geben Auskunft über das Verhältnis der IP. zu den einzelnen ärztlich feststellbaren Schwachsinnsgruppen, den Debilen, Imbezillen, Idioten, sowie den moralisch Defekten. In zwei Hinsichten sind diese Gruppen mit Hilfe von IP. meßbar und vergleichbar: so lange sich die I. der Kinder noch in der Entwicklung befindet, gibt der IQ. ein Bild des Hemmungsgrades, dem diese Entwicklung unterliegt; ist ihre geistige Entwicklung im wesentlichen zum Abschluß gekommen, so ist das IA. charakteristisch, bei welchem der Stillstand erfolgt.

Chotzen (53) hat über 200 Hilfsschüler, die zum größeren Teil dem Alter von 8 und 9 Jahren, zum kleineren Teil höheren Jahrgängen angehörten, sowohl psychologisch geprüft wie ärztlich untersucht und auf Grund der Diagnose in die vier Gruppen: Nichtschwachsinnige (33), fraglich Debile (29), Debile (55), Imbezille (111) eingeordnet. Er gab das Prüfungsergebnis nur als

Differenzwert zwischen LA. und IA. an und fand, daß dieser IR. im Durchschnitt — wie selbstverständlich — von Gruppe zu Gruppe wuchs, daß er aber auch innerhalb jeder Gruppe mit steigendem Alter stark zunahm. In diese verwickelten Verhältnisse brachte nun die von mir vorgenommene Umrechnung der Ergebnisse in IQ. eine gewisse Klärung; Tab. XIII enthält diese Werte. (Diejenigen Zahlen, die wegen zu geringer Personenzahl nicht als eigentliche Durchschnittsziffern gelten können, sind eingeklammert.)

Tabelle XIII. Durchschnittliche Intelligenz-Quotienten.

		Nicht-Schwachs.	Deb.?	Deb.	Imb.
Lebensalter	8	0,92	0,84	0,76	0,71
	9	0,85	0,81	0,77	0,67
	10	(0,80)	(0,80)	0,74	0,62
	11	(0,73)	(0,68)	0,71	(0,64)
	12	(0,75)	(0,75)	(0,73)	(0,61)
	13		(0,73)		

Die Tabelle zeigt nun für jede der beiden Hauptschwachsinnsformen zwar nicht konstante, aber doch sehr ähnliche IQ. durch mehrere Lebensjahre hindurch. Insbesondere sind die Debilen überraschend gleichförmig; variiert doch ihr Durchschnittsquotient durch alle fünf Jahrgänge von 8—12 nur innerhalb des engen Gebietes von 0,71—0,77; ihre I. ist also, an der normalen gemessen, grob ausgedrückt, eine „Dreiviertelsintelligenz". Die Imbezillen weisen etwas größere Schwankungen auf; doch stimmen die IQ. wenigstens für die Altersstufen 9—11 recht gut überein; sie verleihen wiederum, grob gesprochen, ihren Inhabern eine knappe „Zweidrittelsintelligenz".

Die beiden bei Chotzen zuerst genannten Gruppen sind so spärlich vertreten, daß die Durchschnittswerte höchstens für die niederen Jahrgänge beachtet werden dürfen; und hier stimmt der IQ. trefflich mit der ärztlichen Diagnose überein. Die „Nichtschwachsinnigen" haben einen IQ. nahe an 0,9 und die Fraglich-Debilen bilden mit ihrem IQ. zwischen 0,80 und 0,84 eine tatsächliche Zwischenstufe zwischen den Nichtschwachsinnigen und den eigentlich Debilen. — Die vereinzelten Fälle älterer Kinder (im ganzen 7), die Chotzen diesen beiden Gruppen zuteilte, würden ihrem IQ. nach zum größten Teil den Debilen zuzurechnen sein. Vielleicht kann in solchen Fällen der IQ. die schwankende medizinische Diagnose unterstützen.

Da den obigen Quotientenlisten der Vorwurf gemacht werden könnte, daß sie nur Durchschnittswerte enthalten, und daß diese zum Teil auf zu geringen Anzahlen von Einzelwerten aufgebaut sind, habe ich noch eine Berechnung aufgestellt, in der die IQ. der einzelnen Kinder festgestellt und in ihrer Häufigkeitsverteilung registriert wurden. Hierbei wurde vom Lebensalter abgesehen. Es wurden bei der Zählung immer die Werte innerhalb eines Zehners zusammengefaßt, also die IQ. zwischen 0,91 und 1,00, zwischen 0,81 und 0,90 usw.

Tabelle XIV enthält die so gewonnene Verteilung für jede Schwachsinnsgruppe, in absoluten und in Prozentzahlen; die umstehende Figur zeigt zugleich eine graphische Darstellung der prozentuellen Häufigkeiten.

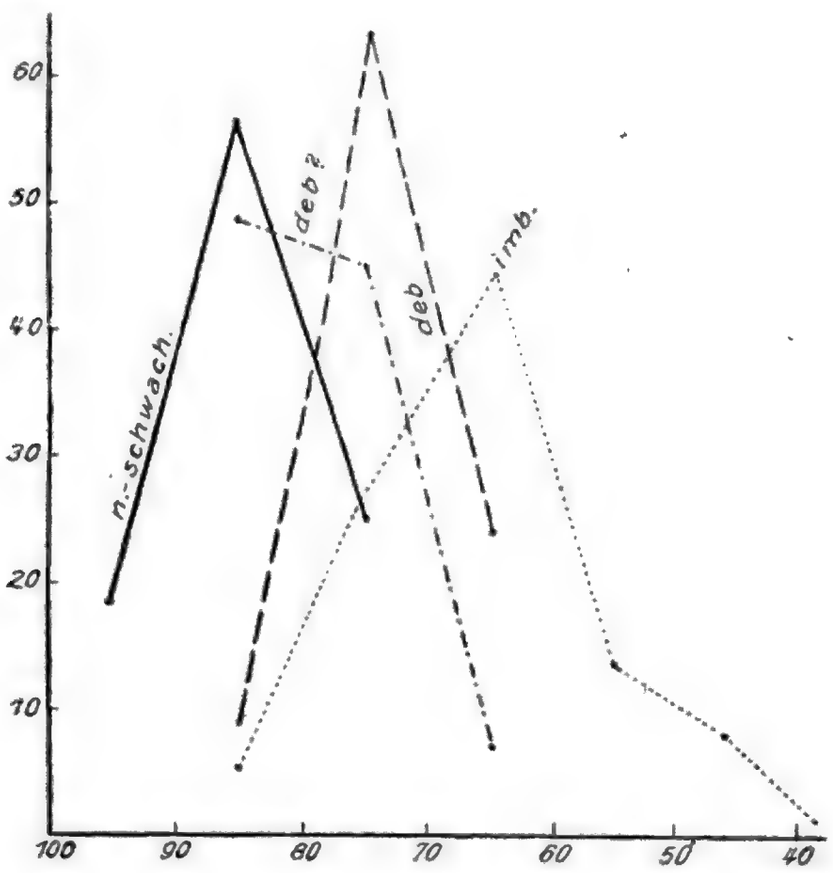
Tabelle XIV. Häufigkeiten der IQ. in den verschiedenen Schwachsinnsgruppen.

IQ.	nicht-schwachs. Häufigkeit		deb.? Häufigkeit		debil Häufigkeit		imbez. Häufigkeit	
	abs.	%	abs.	%	abs.	%	abs.	%
von 0,91—1,00	6	18						
„ 0,81—0,90	19	57	14	48	5	9	6	5½
„ 0,71—0,80	8	25	3	45	37	67	30	27
„ 0,61—0,70			2	7	13	24	49	44
„ 0,51—0,60							15	13½
„ 0,41—0,50							9	8
„ 0,31—0,40							2	2
Insgesamt	33	100	29	100	55	100	111	100

Es ergibt sich eine deutliche Scheidung der Häufigkeitsmaxima für die Hauptgruppen, und zwar haben — ganz entsprechend unserem früheren Ergebnis — die Nichtschwachsinnigen am häufigsten IQ. zwischen 0,81 und 0,90, die Debilen zwischen 0,71 und 0,80, die Imbezillen zwischen 0,61 und 0,70. Bei den Imbezillen ist die Streuung der IQ. größer als bei den anderen Gruppen, wie die Durchschnittswerte auch schon gezeigt hatten. Bemerkenswert ist die ziemlich symmetrische Form der drei Kurven; sie bewirkt, daß sich Häufigkeitsmaximum und Durchschnittswert einigermaßen decken.

Der Übergangscharakter der Gruppe „fraglich debil" kommt ebenfalls zum typischen Ausdruck: ihre Angehörigen verteilen sich ziemlich gleichmäßig auf die Gebiete, die einerseits für die Nichtschwachsinnigen, andererseits für die ausgesprochen Debilen die charakteristischsten sind.

Die Zahl der von Chotzen untersuchten Kinder ist noch nicht groß genug, und vor allem ihre Verteilung auf die verschiedenen Altersstufen nicht weit genug, um obiger Berechnung schon einen endgültigen Wert zuzuschreiben. Neuere Untersuchungen, wie z. B. die von Sauerbrey (s. S. 175) und die von Weigl weisen in die gleiche Richtung; aber eine Nachprüfung mit systematischer Verbindung der psychologischen und ärztlichen Diagnose wäre doch sehr erwünscht.



Als Maß des Schwachsinnsgrades kann der IQ. natürlich nur so lange gelten, als die I. noch nicht zum Stillstand gekommen ist; denn er besagt ja, daß in derselben Zeit, in der die normale I. um ein I.-Jahr wächst, die des Schwachsinnigen nur um dreiviertel oder zweidrittel I.-Jahr zunimmt. Während der Entwicklung bleibt eben die I. des Schwachsinnigen im gleichen Verhältnis hinter der des Normalen zurück. Das hört auf, sobald die I. des Schwachsinnigen zum Stillstand gekommen ist. Denn nun nimmt in der Formel des $IQ = \frac{IA}{LA}$ der Nenner (das Lebensalter) weiter zu, während der Zähler nicht mehr wächst, und zwar geschieht dies schon auf Altersstufen, auf denen der Normale sich noch weiter

entwickelt. Der Schwachsinnige hat eben nicht nur ein langsameres Entwicklungstempo und Ausmaß als der Normale, sondern ist auch früher als dieser am Ende seiner Entwicklung angelangt; und so wird in dieser Zeit die Kluft zwischen beiden gewaltig erweitert.

Aus diesem und manchem anderen Grunde ist es wichtig, neben dem IQ. der einzelnen Schwachsinnsgruppen auch ihre Entwicklungsgrenze festzustellen, d. h. dasjenige IA., das, einmal erreicht, einen endgültigen Schlußpunkt für die Prüfungsleistung bedeutet, so daß auch das zunehmende Lebensalter nichts mehr daran ändert. Die These, die Binet für diese Maße des „I.-Stillstandes" aufstellte (s. S. 169), scheint nach weiteren Untersuchungen im wesentlichen zutreffend zu sein.

Goddard (*143*) hat, wie bereits erwähnt, 400 Zöglinge einer großen Schwachsinnigenanstalt geprüft und ohne Rücksicht auf das Lebensalter lediglich nach dem IA. geordnet. Sie verteilten sich ziemlich gleichmäßig mit je 10—11% auf die IA.-Staffeln von 1—9 Jahren, während die Staffeln 10—12 Jahre zusammen nur mit 7% vorkamen. Wenn Goddard auch hierbei bedauerlicherweise diejenigen Kinder, die ihrer Jugend wegen vielleicht noch über ihre augenblickliche Staffel hinauskommen konnten, und die anderen Insassen, deren Entwicklung längst abgeschlossen war, zusammenwarf, so bleibt doch bestehen, daß Schwachsinnige das IA. 9 nur ganz ausnahmsweise überschreiten.

Durch Vergleiche dieser IA. mit den ärztlichen Diagnosen ergab sich für Goddard folgende Aufstellung:

Es erreichen:	Idioten	stark	mittel Imbezille	schwach	debile („morons")
das IA. von	1 2	3 4	5	6 7	8 9 10 11 12 (10 11 12: selten)

Weitere amerikanische sowie auch die deutschen Untersuchungen decken sich im wesentlichen mit diesen Befunden[1]). So fanden Bloch und Lippa (*128*) bei wiederholter Prüfung von Kattowitzer Hilfsschulkindern, daß diejenigen Prüflinge, welche bereits bei der ersten Prüfung die IA. 9 und 10 erreicht hatten, ein Jahr später fast durchweg keinen Fortschritt zeigten; sie hatten eben schon die Grenze des für sie überhaupt zu Bewältigenden

[1]) Vgl. z. B. Kramer (*155*), S. 29. Chotzen (*53*), S. 494. Sauerbrey (*163*).

erreicht. Die anderen Kinder mit tieferem IA. waren dagegen ein Jahr später meist auf eine höhere Stufe gelangt.

Vor einer falschen Auffassung des Begriffes „Stillstand" möge noch ausdrücklich gewarnt werden. Ein Imbeziller, der Zeit seines Lebens nicht über das „IA. 7" hinauskommt, ist darum nicht einfach einem siebenjährigen Kinde gleichzusetzen; er wächst in vielen Punkten über jenen Geisteszustand hinaus, er sammelt Erfahrungen, die ein normales siebenjähriges Kind nicht hat, erlernt manche Fertigkeiten, erlebt in sich das Erwachen von Trieben und Bedürfnissen, die eben das höhere Alter mit sich bringt. Der Stillstand bezieht sich also nur auf jene Gruppe spezifischer I.-Leistungen, die in den Tests geprüft werden. Auch von diesen mag noch die eine oder andere eine Entwicklung zeigen (vgl. dazu S. 187), aber es bleiben doch so viele endgültige Defekte übrig, daß im Gesamtergebnis das Erreichen eines höheren Wertes als IA. 7 unmöglich wird.

Die Unbedingtheit dieses „Stillstandes" tritt besonders bei Prüfung erwachsener Schwachsinniger hervor. Miß Hinkley (*148*) prüfte vergleichend die jugendlichen und die erwachsenen Insassen einer großen amerikanischen Schwachsinnigenanstalt mit dem Ergebnis, daß die jüngeren (13- bis 14jährigen) Prüflinge durchschnittlich auf dem IA. 9 standen und damit um $^1/_2$ Intelligenzjahr die Durchschnittshöhe der erwachsenen Schwachsinnigen überragten. Die Verfasserin sieht den Grund darin, daß die Älteren eine größere Trägheit und Ungewandtheit im Beantworten von Fragen besitzen und daher insbesondere in jenen Tests rückständig seien, die eine dem Alltag unbekannte Forderung stellen. (Aber mag nicht vielleicht auch mitsprechen, daß die leichteren Schwachsinnsgrade beim Erwachsenwerden die Anstalt verlassen, um doch einen Beruf zu ergreifen, so daß die älteren Jahrgänge sich nur aus schwereren Fällen zusammensetzen?) Bei den Erwachsenen war die „Staffelstreuung" besonders groß; die Binetsche Anordnung der Schwierigkeitsabstufung der Tests traf für sie in weit geringerem Grade zu als für die jüngeren.

4. Das Verhalten zu den einzelnen Tests.

Mit der Möglichkeit der Graduierung ist die Bedeutung der BS.-Methode für die Untersuchung Schwachsinniger keineswegs erschöpft. Vielleicht noch wichtiger als diese ist die mit Hilfe der Methode mögliche qualitative Analyse des einzelnen Prüflings und die Feststellung, wie die einzelnen Tests an den Resultantenwerten beteiligt sind. Die Untersuchung von Chotzen, der zum ersten Male an diese Probleme heranging, hat bewiesen, welche verwirrende Mannigfaltigkeit von Gesichtspunkten hier auftaucht.

Zunächst muß sich ja gleich die Frage aufdrängen: Haben wir denn überhaupt ein Recht dazu, ein 10jähriges debiles Kind, weil die Testberechnung das IA. 7 ergibt, einem 7jährigen normalen Kinde gleichzusetzen, also den Schwachsinn als ein wirkliches „Zurückbleiben" zu kennzeichnen? Es wird ja oft genug behauptet, daß dieser Ausdruck irreleitend sei, da Schwachsinn etwas qualitativ anderes als Normalität sei. Nur die BS.-Methode ermöglicht es, diese Vergleichung exakt durchzuführen.

Und da erkennt man denn in der Tat, daß das IA. 7 beim 10jährigen debilen Kind nicht auf ganz dieselbe Weise zustande kommt wie das gleiche IA. beim 7jährigen normalen Kind; die Staffelstreuung ist nämlich beim schwachsinnigen Kinde sehr viel größer als beim normalen. Das heißt also: die Treffer und Versager verteilen sich beim älteren schwachsinnigen Kinde über sehr viel mehr Altersstaffeln der Tests als beim jüngeren normalen; der Debile versagt unerwartet bei einzelnen ganz leichten Tests, er genügt dafür vereinzelten viel höheren Anforderungen. Es ist eine gewisse Dissoziation zwischen Leistungsfähigkeiten eingetreten, die normalerweise in stärkerer Korrelation stehen.

Wir sind nun auch in der Lage, einen Hauptgrundsatz für diese Dissoziation ausfindig zu machen. Es gibt nämlich gewisse Leistungen, die wesentlich eine Funktion des Alters sind, relativ unabhängig von der I., andere, die rein durch eine bestimmte Höhe der erreichbaren I.-Entwicklung bestimmt sind, gleichgültig, in welchem Alter diese erreicht wird. In bezug auf die erste Gruppe wird das 9- und 10jährige Kind, auch wenn es debil ist, weiter sein, als das 6- und 7jährige normale Kind; bei der zweiten Gruppe von Leistungen muß der jüngere Normale den älteren Schwachsinnigen übertreffen[1]).

A priori wird man vermuten, daß zur Gruppe der Leistungen mit reiner Altersentwicklung solche gehören werden, die von einer Summe häufig gemachter Alltagserfahrungen und ständig geübter Alltagsverrichtungen abhängen. Aber solche vagen Vermutungen nützen nicht allzu viel, und so wird es denn von Wert sein, daß wir aus der Analyse der BS.-Ergebnisse lernen können, welche Tests bei den Schwachsinnigen mehr mit dem Alter und welche mehr mit der eigentlichen Intelligenz korrelieren. Bisher sind hierfür haupt-

[1]) Vgl. hierzu den Abschnitt über Begabungs- und Entwicklungstests, S. 151 ff.

sächlich die Befunde von Chotzen verwendbar; doch geben sie nur ein unvollständiges Bild, weil Ch. es ganz überwiegend mit Schwachsinnigen einer bestimmten Altersgruppe (8—9jährigen) zu tun hatte.

Ch. gibt eine ganze Reihe verschiedener Berechnungen, um die „Wertigkeit" der verschiedenen Tests für die Schwachsinnsdiagnose nachzuweisen. Am einfachsten ist natürlich der direkte Vergleich von schwachsinnigen mit normalen Kindern gleichen Lebensalters (nach Bobertags Befunden).

Hier war die Rückständigkeit der Schwachsinnigen am geringsten bei folgenden Tests[1]): Vor- und Nachmittag unterscheiden; Definitionen durch Zweckangaben; Altersangabe; ästhetisches Urteil; Angabe der Fingerzahl; Bildbeschreibung; Abzählen von 13 Pfennigen. Dagegen war die Rückständigkeit sehr groß bei: Nachsprechen von 16 Silben und von 5 Zahlen; auf 1 Mark 80 Pfennige herausgeben; Rückwärtszählen von 20 bis 0; Definition durch Oberbegriff; Vergleichung zweier Gegenstände aus dem Gedächtnis; Wiedergabe einer kleinen Geschichte; Monate aufzählen; 5 Gewichte ordnen.

Bei einer anderen Altersstufe wird sich der Eindruck vermutlich wieder ändern. So ist es zweifellos, daß die Bilderklärung, die von älteren verlangt wird, einen gewaltigen Unterschied zwischen Schwachsinnigen und Normalen zeitigen würde, während die von jüngeren geforderte Bildbeschreibung nach Chotzen keine solche Wirkung hat.

Immerhin zeigt auch schon obige Aufzählung Chotzens, daß die geringe Verschiedenheit namentlich bei solchen Tests besteht, die sich auf häufig geübte Leistungen (Abzählen, Altersangabe) und Alltagserfahrungen (Fingerzahl, Vor- und Nachmittag) beziehen; dagegen springt die ganze Rückständigkeit des Schwachsinnigen sofort hervor, wo Ungewohntes verlangt, Neues dargeboten und eine scharfe Aufmerksamkeitsspannung beansprucht wird.

Einen ähnlichen Vergleich kann man nun aber innerhalb der Hilfsschüler vornehmen zwischen den verschiedenen Schwachsinnsgruppen, wie sie die ärztliche Diagnose zeigt; Ch. stellte fest, welche Tests bei gleichem Lebensalter von einer Gruppe zur anderen besonders stark abnahmen. Ich erwähne nur diejenigen, welche deutlich doppelten Abfall zeigten (von den Nichtschwachsinnigen zu den Debilen, und dann wiederum zu den Imbezillen)[2]).

---

[1]) *53*, S. 440.
[2]) *53*. S. 458.

Bei 8- bis 9jährigen Kindern: Rhombus abzeichnen, 5 Zahlen nachsprechen, leichte I.-Fragen; etwas geringer war der Abfall für: fünf Münzen, Vergleich von zwei Gegenständen.

Bei älteren Kindern (Ch. hatte nachträglich zu Vergleichszwecken noch eine Reihe älterer Hilfsschüler geprüft): Vergleich, Wiedergabe einer Zeitungsnotiz, 5 Gewichte ordnen, auf 1 Mark 80 Pfennig herausgeben, Definitionen durch Oberbegriffe, Kenntnis aller Münzen, Nachsprechen von 5 Zahlen.

Eine dritte Berechnung veranstaltet endlich Ch. durch Vergleichung derjenigen Prüflinge, die auf demselben IA., aber in verschiedenem LA. stehen. Dasselbe IA. (z. B. 6 Jahr) kann dadurch zustande kommen, daß ein 8jähriges Kind um 2, oder daß ein 9jähriges um 3, oder daß ein 10jähriges um 4 Jahr im Rückstand ist. Nun gibt es Tests, welche bei Kindern einer I.-Stufe deutliche Zunahme der Leistungen mit dem Lebensalter zeigen, andere, die Konstanz, endlich solche, die sogar eine Abnahme aufweisen. Die erstgenannten Tests mit „Alterszuwachs" sind zweifellos solche, die am wenigsten mit der I. zu tun haben, da sie bei gleicher I. von den älteren Kindern besser geleistet werden. Dagegen stehen die anderen Tests deutlich mit der I. in Korrelation, ganz besonders diejenigen, bei denen die älteren Kinder sogar schlechter abschneiden. Es ergibt sich nun:

Starken Alterszuwachs zeigen: Abschreiben, Diktat schreiben, das Wiedergeben von zwei Einzelheiten aus einer Geschichte, Aufzählen der Wochentage.

„Die Tests mit dem starken Alterszuwachs betreffen also ausschließlich Kenntnisse, insbesondere Schulkenntnisse, bei deren Aneignung es auf die Dauer des Unterrichts ankommt. Wo nur ein geringer Zuwachs zu verzeichnen ist, da spielen Kenntnisse bei einigen wohl auch eine Rolle (5 Münzen, Alter), zumeist aber sind es Tests, bei denen Übung, aber auch die natürliche Zunahme der Leistungsfähigkeit die Resultate fördern wird (z. B. Ausführen dreier Aufträge, Rückwärtszählen, Nachsprechen von 16 Silben). Bei allen diesen ist der Alterszuwachs gering. Ganz ohne ihn sind Urteils- und Kombinationsleistungen oder solche, die hohe Anforderungen an die Auffassung stellen (Vergleiche, I.-Fragen, Lücken erkennen, 5 Zahlen nachsprechen)[1])."

Wahrscheinlich gehören zur letzten Gruppe auch: Wiedergabe von sechs Einzelheiten aus einer Geschichte, Ordnen der 5 Gewichte, Bilderklärung, 80 Pfennig auf 1 Mark herausgeben: doch sind hier die Zahlen für feste Ergebnisse zu klein.

Vergleicht man nun diese auf verschiedenem Wege gewonnenen Listen miteinander, so zeigen sie zwar in manchen Einzelheiten Abweichungen; aber im großen und ganzen treten doch immer wieder dieselben Tests auf, als diejenigen, in welchen sich der I.-Defekt unverhüllt und unkompensiert offenbart, während bei

[1]) *53*, S. 453.

anderen die verringerte I. durch das höhere Alter wettgemacht werden kann.

Zur Ergänzung obiger Angaben sei auf die freilich recht dürftigen Angaben hingewiesen, die Bloch und Lippa (*128*) auf Grund der Prüfung von 71 zweimal geprüften Hilfsschulkindern verschiedenen Alters machen. Als besonders kennzeichnend für Schwachsinnige haben nach ihnen solche Tests zu gelten, bei denen erstens die Leistung gegenüber normalen Kindern einen besonders starken Rückstand zeigt, und bei denen zweitens die wiederholte Prüfung nach einem Jahr keinen Altersfortschritt aufweist. Solche Tests sind z. B.: das Ordnen von 5 Gewichten, das Nachsprechen von mehr als drei Ziffern, das Herausgeben von 80 Pfennig auf eine Mark, Aufsagen der Wochentage, Wiedergabe von Hauptpunkten einer vorgelesenen Geschichte.

Auch aus den Weiglschen Tabellen (*174*) lassen sich die Tests herausziehen, in denen die Hilfsschulkinder besonders weit zurückstehen: es sind zum großen Teil dieselben, die Chotzen und Bloch-Lippa fanden: Behalten von 5 Ziffern, Wiedergabe der Zeitungsgeschichte, 5 Gewichte ordnen, ferner die höheren Formen der Bildbeschreibung (Erkennung der Pointe) und der Definition (durch Oberbegriffe).

Eine Fortsetzung solcher Analysen bei größeren Zahlen von Schwachsinnigen aus verschiedenen Lebensaltern wäre dringend zu wünschen.

5. Intelligenzprüfungen sittlich Minderwertiger.

Besonderes Interesse verdienen IP. bei solchen Jugendlichen, die sich als moralisch minderwertig erwiesen haben. Gelegenheiten zu ihrer Untersuchung bieten einerseits die Jugendgerichte, andererseits Schwachsinnigen- und Fürsorgeanstalten. Es zeigte sich nun durchweg, daß starker sittlicher Defekt nicht mit gleich starkem intellektuellem Defekt verbunden ist. Die moralische Haltlosigkeit kann in den Formen der Ungehemmtheit des Trieblebens, der Willensschwäche und der Gefühlsstumpfheit als offenkundig pathologische Erscheinung auftreten, während die I. nur wenig — oder überhaupt nicht merklich — im Rückstand zu sein braucht.

Belege hierfür lieferten zunächst die Untersuchungen von Goddard (*143*). In der Schwachsinnigenanstalt zu Vineland befanden sich 22 Individuen mit moralischem Schwachsinn; die IP. ergab für diese, daß 15 das IA. 9, 5 das IA. 10 und je 1 das

IA. 11 und 12 hatten, daß demnach ihr geistiger Stand nie unter leichte Debilität herabgegangen war. Jugendliche Delinquenten untersuchten in Amerika Williams (*177*), in Deutschland Kramer (*156*). Unter den 150 Prüflingen des Erstgenannten waren 22% knapp normal, 25% Grenzfälle und 28% schwachsinnig, und zwar durchweg debil. Kramer berichtet über die Prüfung Jugendlicher, die ihm von der Jugendfürsorge zur Begutachtung überwiesen worden waren, folgendes: Bei den (kriminellen und verwahrlosten) Jugendlichen, die meist im Alter zwischen 12 und 16 Jahren standen, fehlten schwerere intellektuelle Defektzustände fast ganz. Weitaus vorherrschend waren die IA. 9—11. Von denjenigen 23 Prüflingen, bei welchen eine offenkundige moralische Psychopathie vorhanden war, waren 11 nach der IP. durchaus normal (von denen manche durch ihre psychomotorische Regsamkeit sogar übernormale I. vortäuschten); bei den übrigen entsprach der intellektuelle Defekt durchaus nicht der Schwere des moralischen[1]).

Wie ist nun diese Tatsache zu erklären? Goddard stellt für sein Ergebnis, daß die moralisch Minderwertigen nicht unter das IA. 9 herunterreichen, folgende Vermutung auf: Gewisse unsittliche Instinkte, wie Trieb zum Lügen, zum Stehlen usw. erwachen normalerweise um das neunte Jahr herum; später entwickelt sich die Vernunft, welche die Hemmung diesen Instinkten gegenüber liefert. Bei Kindern unter dem IA. 9 sind jene Instinkte noch nicht entwickelt, die Kinder aber, die ungefähr auf IA. 9 stehenbleiben, zeigen jene Instinkte, ohne bis zu der Entwicklung der Hemmung zu gelangen und werden so sittlich defekt.

Dieser Erklärungsversuch ist wohl etwas eng gefaßt. Doch scheint mir so viel daran richtig zu sein, daß sittliche Minderwertigkeit als pathologische und zugleich antisoziale Erscheinung nur dort möglich ist, wo weder eine zu geringe noch eine zu hohe I. vorhanden ist. Menschen mit stark rückständiger I. (Imbezille und Idioten) mögen wohl in moralischer Hinsicht völlig passiv

[1]) Über einen entgegengesetzten Befund berichtet der Ungar Bela Tabajdi Kun (*126*a), der bei jugendlichen Kriminellen weit schwerere I.-Defekte als bei Hilfsschülern gefunden haben will. Die Arbeit ist aber lediglich ein Beleg für den Mißbrauch, der mit der BS.-Methode getrieben werden kann. Denn die geprüften Hilfsschüler standen zum Teil, die jugendlichen Kriminellen sämtlich jenseits der Altersgrenze, für die die Tests der BS.-Methode ausreichen und IA.-Berechnungen zulassen — kein Wunder, daß die 14—19jährigen Kriminellen nach dieser Methode besonders rückständig erscheinen mußten!

und gleichgültig sein, Willensschwäche und Gefühlsstumpfheit zeigen; aber zur Umsetzung dieser Amoralität in tätige Antimoralität fehlt ihnen jene Einsicht und Urteilsfähigkeit, die zur Durchführung krimineller Handlungen erforderlich ist. Solche Individuen werden mehr zu den sozial Unbrauchbaren, als zu den sozial Gefährlichen gehören. Andererseits wird eine hohe I. zwar niemals allein einen sittlich guten Willen gewährleisten (es gibt sehr kluge Verbrecher); aber sie wird doch immerhin diejenigen Denkmittel zur Verfügung stellen, die als Voraussicht, Abwägung der Vorteile und Nachteile, Einsicht in die Verbindlichkeit sittlicher, staatlicher und religiöser Forderungen usw. eine Hemmungswirkung auf das reine Triebleben auszuüben vermögen.

Als Gesamtergebnis bleibt bestehen, daß sittliche Minderwertigkeit in relativer Unabhängigkeit vom I.-Grad bestehen kann und daß insofern die alte Bezeichnung „moral insanity" nicht ganz unrichtig gewesen ist.

Während der Schulzeit drückt sich moralische Rückständigkeit zum größten Teil in der Unfähigkeit aus, sich den erziehlichen Anforderungen der Schule zu fügen; und auch hier zeigt sich wieder eine relative Unabhängigkeit dieser „Schul-Unfähigkeit" von der I. So gibt es unter den 8- und 9jährigen Prüflingen Chotzens eine große Zahl, die nicht den (von Binet für die Zuweisung zur Hilfsschule geforderten) Rückstand von 2 I.-Jahren hatten und dennoch durchaus in die Hilfsschule gehörten, weil sie in der Normalschule völlig versagten. Ebenso fand Kramer bei den von der Jugendfürsorge überwiesenen älteren Jugendlichen folgende Ziffern: Während in 39 Fällen I.-Grad und Schulalter des Prüflings übereinstimmten, stand in 36 Fällen die Schulfähigkeit hinter der I. mehr oder weniger erheblich zurück.

Es zeigt sich somit bei den abnormen Individuen eine Erscheinung, die im folgenden Kapitel auch bei den normalen nachzuweisen sein wird: daß die Schultüchtigkeit nicht allein durch Eigenschaften der geistigen Begabung bestimmt wird. Zum Teil beruht ja ein solcher, nicht intellektuell bedingter Schulrückstand auf äußeren Umständen: schlechten häuslichen Verhältnissen, Verwahrlosung, Orts- und Schulwechsel, längerer Krankheit. Zu einem anderen Teil wirkt hier aber ein inneres Manko mit: jene Willenseigenschaften, welche zur I. ergänzend hinzutreten müssen, um brauchbare Menschen zu erzeugen, sind nicht der I. entsprechend entwickelt. „Kinder dieses Typus fügen sich naturgemäß sehr

schlecht in die Schuldisziplin, schwänzen den Unterricht, treiben allerlei Unfug in der Klasse, sind durch Strafen ganz unbeeinflußbar usw., so daß sie trotz guter I. mehr oder minder häufig sitzenbleiben. Als besonders ungünstig erwiesen sich diejenigen Fälle, in denen neben diesen psychischen Anomalien auch intellektuelle Defekte geringen Grades bestanden" (Kramer (*155*), S. 31).

Anhangsweise sei auf drei weitere Untersuchungen an moralisch abwegigen Jugendlichen hingewiesen, bei der nicht die B.-S.-Methode, sondern andere Prüfmittel der I. verwertet wurden. Die Ergebnisse stimmen gut zu den bisher genannten.

Margot Dósai Révész (*138*) prüfte 40 Knaben einer ungarischen Fürsorgeerziehungsanstalt auf drei Funktionen: Rechenfähigkeit, Gedächtnis (für Wortpaare), Aussagefähigkeit. Wurden nun die Prüflinge in die Gruppen der geistig normalen, der moralisch Verkommenen und der intellektuell Schwachsinnigen gegliedert, so ergab sich, daß die Testleistungen fast durchgehends die eben angegebene Reihenfolge zeigten, d. h. die moralisch Defekten standen zwar hinter den normalen Intelligenzen zurück, übertrafen aber meist die Knaben mit ausgesprochenem I.-Defekt.

Riebesell (*161*a) untersuchte 125 Insassen aus Hamburger Waisenhäusern und Erziehungsanstalten für Verwahrloste mit einer Reihe von Tests, die zum Teil denen von Moede und Piorkowski nachgebildet waren (Ausstreichen bestimmter Buchstaben, Lückenergänzung, Dreiwortmethode usw.). Er fand bei den Verwahrlosten keinen I.-Rückstand gegenüber den moralisch-normalen Waisenhauszöglingen.

Weit eingehender noch sind die Analysen, die Gregor und Voigtländer (*145*, S. 162 u. 144) an je 100 männlichen und weiblichen Verwahrlosten, Insassen einer Heilerziehungsanstalt bei Leipzig, vornahmen. Diese Jugendlichen wurden zunächst nach ihrer moralischen und kriminellen Beschaffenheit geschildert; sodann wird ihre I. einerseits auf Grund der klinischen Beobachtung beurteilt, andererseits mittels der Definitionsmethode geprüft. Zwei Ergebnisse sind für uns von Wichtigkeit. Die Art, wie sich die Prüflinge zu der Aufgabe stellten, bekannte Begriffe zu erklären, ist nach Gregor und Voigtländer sehr symptomatisch für ihre I., ja darüber hinaus auch für ihre charakterologische Beschaffenheit. Freilich kann man sich nicht ganz des Eindrucks erwehren, daß hier eine Überschätzung eines immerhin einseitigen I.-Tests vorliegt (vgl. S. 107 dieses Buches). Sodann tritt auch hier die teilweise Unabhängigkeit des moralischen Defekts von intellektueller Rückständigkeit hervor. Insbesondere sind sehr viele „asoziale" Jugendliche von ausgesprochen unversehrter I., während man unter den „moralisch schwachen" und „moralisch minderwertigen" Individuen zu etwa gleichen Teilen geistige Normalität und geistige Beschränktheit bis Minderwertigkeit findet. Höhere Grade geistiger Unternormalität (Debilität und Imbezillität) sind unter den Verwahrlosten sehr selten, ein Ergebnis, das sich durchaus mit den oben angegebenen Befunden anderer Forscher deckt.

Alle im obigen besprochenen Untersuchungen von moralisch Defekten arbeiteten mit I.-Tests, die mit Moralfragen nichts zu tun hatten. Ein anderes Problem tut sich auf, wenn wir diejenige Seite der I., die sich auf Moralisches

bezieht, gesondert untersuchen, also die moralische Einsicht, die moralische Urteils- und Kritikfähigkeit. Es sind hierfür neuerdings Prüfmethoden ausgebildet worden, die fälschlich Methoden der „Gesinnungsprüfung" genannt werden; über sie ist an anderer Stelle (S. 98ff.) berichtet. Die Anwendung dieser Methoden an sittlich defekten Jugendlichen wird wohl deutlichere Ausfälle als die indifferente IP. ergeben; doch sind exakt vergleichende Untersuchungen meines Wissens noch nicht in der Öffentlichkeit bekannt geworden.

X. Kapitel.

Intelligenz und Schulleistung.

Das Verhältnis der I. zur Schulleistung (SchL.) ist von unmittelbarer Wichtigkeit für Fragen der praktischen Pädagogik; denn hier dürfen wir hoffen, einen Einblick in die Faktoren zu gewinnen, welche das Fortkommen der Kinder in der Schule, den Platz, den sie auf Grund ihrer Leistungen unter ihren Mitschülern einnehmen, den Ausfall ihrer Zeugnisse bedingen. Man ist im allgemeinen geneigt, die Beziehung dieser schulischen Leistungsfähigkeit zur intellektuellen Leistungsfähigkeit für eine sehr enge zu halten: gute Schüler werden von vornherein als intelligent angesehen, und von intelligenten Kindern werden mit einer gewissen Selbstverständlichkeit gute Schulleistungen erwartet, umgekehrt bei den schlechten Gruppen. Solange wir freilich keine besonderen Prüfungsmittel der I. hatten, fehlten alle Grundlagen für eine exaktere Kenntnis dieser Verhältnisse; man mußte sich mit Vermutungen oder mit Verallgemeinerungen gelegentlicher Beobachtungen begnügen.

Das ist nunmehr anders geworden; wir sind imstande, die mit exakten Methoden festgestellte I. mit der SchL. in Korrelation zu setzen, und es ergibt sich, daß die Beziehungen zwischen beiden Faktoren durchaus nicht so eng und eindeutig sind, wie man früher meist geglaubt hatte.

Zum Zweck einer solchen Vergleichung müssen sowohl für die I. wie für die SchL. geeignete Maßstäbe gewählt werden.

Ein gröberes Verfahren begnügt sich hier mit der Herstellung einiger weniger Gruppen auf jedem der beiden Gebiete: so wird zur Bestimmung der SchL. das Klassenalter oder die Gesamtzensur gewählt und mit dem I.-Alter nach dem Binetschen Staffelverfahren verglichen. Die Ergebnisse solcher gröberen Vergleichung werden uns im ersten Abschnitt dieses Kapitels beschäftigen.

Aber diese Maßstäbe reichen für feinere Feststellungen nicht mehr aus, insbesondere dann nicht, wenn man I.- und SchL.-Unterschiede innerhalb einer ziemlich homogenen Gruppe, also z. B. bei Kindern gleichen Alters oder gleicher Schulklassen, untersuchen und in gegenseitige Beziehung setzen will. Hier muß man versuchen, an Stelle der allzu groben Staffelprüfung Untersuchungsmittel der I. zu wählen, die für feinere Unterschiede empfindlich sind; und man muß andererseits die SchL. in weiterem Maße differenzieren, als es etwa durch die 3 oder 5 Zeugnisprädikate möglich ist. Diese Forderung wurde erfüllt durch die Ausbildung von Rangmethoden. Die Angehörigen einer homogenen Gruppe werden nach einem bestimmten Wertgesichtspunkt in eine Rangreihe geordnet, in der nun jeder Schüler einen Rangplatz erhält. Legt man an dieselben Personen verschiedene Bewertungsgesichtspunkte an — z. B. einmal die I., das andere Mal die SchL. —, so erhält man zwei (oder mehr) Rangreihen, für die der Grad der Übereinstimmung durch Korrelationsrechnungen[1]) bestimmt werden kann.

Schulleistungsrangordnungen sind allbekannt und in den verschiedensten Formen in Gebrauch. So erhalten die Schüler Platznummern nach dem Ausfall eines Extemporales; der Fachlehrer ordnet sie zum Vierteljahrsschluß nach der Gesamtheit der Leistungen in seinem Fach; endlich werden alle diese Fachreihen verbunden zu einer „Klassenrangordnung", in der jeder Schüler seinen „Klassenplatz" erhält.

Intelligenzrangordnungen sind auf zwei Arten herstellbar: durch Prüfung und durch Schätzung. Gewisse Tests oder Testreihen können so gestaltet werden, daß sie eine mehr oder minder große Anzahl verschiedener Lösungsgrade aufweisen; hiernach wird eine Testrangordnung gebildet. Daneben aber tritt hier als ganz neue Methode die Intelligenzschätzung, die sich nicht auf das Experiment, sondern auf die Beobachtung und Beurteilung des Lehrers stützt. Sie bildet ein außerordentlich wertvolles Mittelglied zwischen der Leistungsreihe und der Testreihe und kann nach beiden Seiten hin zur Kontrolle und Sicherung der Ergebnisse dienen; zugleich ist sie dem Lehrer, der sie anwendet, ein sehr brauchbares Hilfsmittel zum psychologischen Verständnis seiner Klasse.

Indem wir die I.-Schätzung den Prüfungsmethoden als gleich-

[1]) Über deren Methodik s. S. 65 und Anhang.

berechtigt zur Seite stellen, soll der methodische Grundsatz nochmals deutlich zum Ausdruck gebracht werden, daß es nicht genügt, die I. nur in ihren rein reaktiven, künstlich hervorgerufenen Äußerungen festzustellen; die natürliche I. und insbesondere ihre spontane Betätigung verlangt gleiche Rücksicht und muß sogar dazu dienen, die Ergebnisse der Experimentalprüfung zu deuten und zu eichen. Aus diesem Grunde wird die I.-Schätzung vor der Erörterung der experimentellen Rangordnungen besprochen werden.

1. Klassenalter und Schulzeugnisse in ihrer Beziehung zur Intelligenz.

Das Klassenalter ist das normale Alter der Klasse, der das Kind angehört. Nehmen wir das Alter von 6 Jahren als Schulbeginn an, so ist das Klassenalter einer Klasse, die eben in das vierte Schuljahr tritt, $6 + 3 = 9$ Jahre. Befindet sich in dieser Klasse ein Kind von 11 Jahren, so hat es einen Schulrückstand von 2 Jahren, dagegen hat ein 8jähriger Klassengenosse einen Schulvorsprung von 1 Jahre. Schulvorsprünge sind bei uns, wegen des genau geregelten Versetzungsturnus, sehr selten; sie kommen höchstens vor, wenn Kinder nach privater Vorbereitung oder von einer anderen Schule her eingeschult werden. (Im Ausland scheinen sie häufiger zu sein.) Schulrückstände sind dagegen recht häufig infolge des Sitzenbleibens, ferner durch längere Erkrankung usw.; sie gewinnen gelegentlich einen ziemlich hohen Grad.

Die Vergleichungen des Klassenalters mit der Binetschen Staffelprüfung ergab nun eine vorwiegende Übereinstimmung zwischen Rückstand im Klassenalter (d. h. Überalterung) und I.-Rückstand; entsprechend beim Vorsprung auf beiden Gebieten. Die älteren Schüler innerhalb einer Klasse sind gewöhnlich auch die weniger intelligenten; die jüngsten Schüler gehören meist zu den höchsten I.-Stufen. Aber diese Beziehung ist keine durchgehende; es gibt stets Fälle, die sich dieser Regel nicht fügen, indem Kinder, die in der einen Hinsicht rückständig oder vorgeschritten sind, in der anderen zur normalen Mittelgruppe gehören. Ganz selten kommen wohl auch paradoxe Fälle vor: hohe I. bei überalterten Schülern und umgekehrt. Solche erklären sich aber fast immer durch äußere Umstände, wie späte Umschulung vom Dorf in die Stadt, lange Krankheit usw.

Vergleichungen von Klassenalter und IA. sind bisher von Binet, Goddard und Wiersma angestellt worden.

Tabelle XV.

Binet		Die Intelligenz stand unter dem Niveau des Lebensalters	auf dem Niveau des Lebensalters	über dem Niveau des Lebensalters	Summa
Das Klassenalter zeigte	Rückstand	14	9	1	24
	Normalität	16	33	16	65
	Vorsprung	0	5	7	12
	Summa	30	47	24	101

Binet (*4*, S. 162) gibt eine Verteilungstafel von 101 Kindern und findet die Übereinstimmung ziemlich befriedigend. In der Tat fehlen wenigstens paradoxe Fälle so gut wie ganz; im übrigen sind aber doch merkliche Abweichungen festzustellen.

Von den I.-Vorgeschrittenen zeigt nur ein knappes Drittel auch Schulvorsprung, von den I.-Rückständigen ist nicht ganz die Hälfte zugleich schulrückständig; und die schulnormalen Kinder stehen zu einem Viertel über, zu einem Viertel unter der I.-Höhe ihres Lebensalters.

Zu einem entsprechenden Ergebnis kam Goddard (*144*) bei seinem Massenversuch; er fand, daß mehr als die Hälfte aller geprüften Kinder in einer Klasse waren, die nicht ihrem IA. entsprach, und zwar befanden sich die meisten von diesen in einer zu tiefen Klasse, nur wenige in einer zu hohen.

Eine nur teilweise Übereinstimmung zwischen SchL. und IP. fand auch Wiersma (*176*); er hatte Volksschüler von 6—11 Jahren nach BS. geprüft; eine besondere Statistik der sitzengebliebenen (also klassenrückständigen) Prüflinge ergab, daß 71% unter dem I.-Stand, je 14½% dagegen auf und über dem Niveau ihres Alters waren.

Tabelle XVI.

Bobertag		Die Intelligenz stand unter dem Niveau des Lebensalters	auf dem Niveau des Lebensalters	über dem Niveau des Lebensalters	Summa
Zensuren	schlecht	29	17	0	46
	genügend	26	79	21	126
	gut	0	13	31	44
	Summa	55	109	52	216

Den Vergleich des IA. mit den Schulzensuren zog Bobertag (*51*, II, S. 501, Tab. II). Seine Verteilungstafel zeigt wieder das Fehlen der paradoxen Fälle.

Was bedeutet nun diese Unvollständigkeit in der Übereinstimmung zwischen der SchL.-Fähigkeit und dem Ausfall der IP.?

Man könnte einerseits sagen, es sei dies ein Beweis für die Mangelhaftigkeit der Tests. Denn da das Schulalter und das Zeugnis der verdichtete Ausdruck für die ständige und mannigfaltige Leistungsfähigkeit des Kindes und daher weit bezeichnender sei als der Ausfall einer halbstündigen Testprüfung, so sei der letzteren nur dann Vertrauen zu schenken, wenn sie sich mit jener decke; tue sie es nicht, so taugen eben die Tests oder wenigstens die auf ihnen aufgebauten Staffelordnungen nichts.

Nun haben wir ja schon oben die Schwächen der bisherigen Graduierung nach BS. nachgewiesen, und es ist zweifellos, daß diese auf die Unstimmigkeit mit den SchL. mit eingewirkt haben[1]). Denn da z. B. die Tests für 7jährige zu leicht sind, werden auch viele weniger begabte Kinder von 7 Jahren nach der Prüfung noch den Stand ihres Alters erreichen, während sie in der Schule nicht „genügend" stehen. Umgekehrt bei den älteren Kindern. Ich glaube indessen nicht, daß dies die einzige Ursache der Unstimmigkeit sei; der Grund liegt tiefer.

Man kann andererseits meinen, daß nur die Tests ein wirkliches Bild der geistigen Veranlagung geben, und daß daher die ganze Schuld an jener Unstimmigkeit der Schule zugeschrieben werden müsse; die Lehrer hätten eben die Schüler falsch beurteilt, wenn sie ihnen Zeugnisse gegeben hätten, die nicht ihrer I.-Höhe entsprächen, und falsch behandelt, wenn sie sie in einer Klasse zurückhielten, über die sie der I. nach bereits hinausgekommen wären. So schreibt z. B. Goddard jenes Ergebnis fast ganz einer verfehlten Versetzungstechnik auf Rechnung (*144*, S. 241 und 249).

Das heißt aber doch, das Kind mit dem Bade ausschütten. Freilich gibt es auch bei den Lehrern menschliche Irrungen, und nicht wenige ihrer Maßnahmen mögen durch ein falsches Urteil über die geistige Reife des Schülers bedingt sein. Aber daß davon die Hälfte aller Kinder betroffen werden sollte, ist völlig undenkbar.

Mir scheint vielmehr gerade aus jenem Ergebnis hervorzugehen, daß die Voraussetzung der beiden bisher genannten Standpunkte falsch ist. Eine völlige Übereinstimmung von SchL. und IP. ist gar nicht zu erwarten und zu verlangen, weil die SchL. außer von der I. noch von ganz anderen Faktoren abhängt.

---

[1]) Dies wird auch von Bobertag nachgewiesen.

So spielt ja die Stärke des Gedächtnisses, die bekanntlich nur in mäßiger Korrelation zur I. steht, beim Zustandekommen und bei der Beurteilung der Schulleistungen eine große — vielleicht eine zu große — Rolle; auch die verschiedenen Sondertalente kreuzen die Wirkung der allgemeinen I. Sodann aber kommen Faktoren in Betracht, die gar nichts mit der intellektuellen Beschaffenheit zu tun haben, sondern dem Willensgebiet (in weitestem Sinne) angehören: die Intensität und Ausdauer der Aufmerksamkeit, der Fleiß und die Gewissenhaftigkeit, das Pflichtbewußtsein und die Fähigkeit der Einordnung in die soziale Gruppe.

Sie bilden vor allem den Einschlag, der zur I. hinzukommen muß, um die bloße Leistungsmöglichkeit in wirkliches Leisten umzusetzen, und sie sind andererseits imstande, auch da, wo die I. geringer ist, eine achtenswerte Tüchtigkeit zu erzeugen. Das gilt im Leben, und es gilt auch bereits in der Schule; und es ist doch gut, daß diese Verhältnisse einmal mit zahlenmäßiger Eindringlichkeit klargelegt werden konnten. Denn eben dies lehren die Zahlen, daß die I. niemals mehr als ein Teilfaktor der Schultätigkeit ist; und der einseitige Intellektualismus, der an den Schülern nur ihre Begabung beachtet und wertet, darf damit als widerlegt gelten. Darum bleibt doch die beträchtliche Wichtigkeit dieses Faktors bestehen, und es wird seine genauere Kenntnis durch IP. und andere psychologische Hilfsmittel in den Schulen noch viel Nutzen stiften und Irrungen und falsche Behandlung verhindern können. Dies aber um so mehr, je mehr man sich über Umfang und Grenzen seiner Bedeutung klar wird. Wenn jetzt z. B. in einem individuellen Fall gute SchL. mit einem mittelmäßigen Ausfall der IP. zusammentreffen und auch nicht ein Spezialtalent (das ja leicht erkennbar ist) entscheidenden Einfluß gehabt haben kann, so ist eine an Sicherheit grenzende Wahrscheinlichkeit vorhanden, daß die Stärke dieses Schülers vornehmlich in Eigenschaften der Willens- und Charakterseite zu suchen ist.

So ist denn die Unstimmigkeit zwischen IP. und SchL. gerade geeignet, unser Vertrauen zu dem psychologischen Prüfungsverfahren zu erhöhen. Sehr treffend bemerkt hierzu Kramer[1]):

„Wenn wir . . . einen strikten Parallelismus zwischen den Ergebnissen der IP. und den SchL. gefunden hätten, so müßten wir das größte Mißtrauen gegen die Methode haben. Es läge der Verdacht sehr nahe, daß wir nichts weiter tun,

[1]) (*155*), S. 30/31. Bei K. bezieht sich die Bemerkung auf die Prüfung nichtnormaler Kinder; sie gilt aber ebenso für normale.

als die Schulkenntnisse direkt oder auf einem Umwege zu prüfen, und dann wäre die Methode als Begabungsprüfung unbrauchbar und ihre Anwendung überflüssig, da wir uns nur nach den Schulleistungen zu erkundigen brauchten."

2. Intelligenzschätzung.

Unter Intelligenzschätzung (ISch.) verstehen wir die vergleichende Beurteilung und darauf gegründete Herstellung einer Rangordnung der geistigen Allgemeinveranlagungen innerhalb einer Schulklasse auf Grund der dem Lehrer zur Verfügung stehenden Beobachtungen. Die Aufgabe solcher ISch. ist eine doppelte: einmal dem Lehrer zur Selbstbesinnung über die Fähigkeitsverteilung in seiner Klasse und über das Verhältnis der Fähigkeiten zu den Leistungen zu verhelfen, sodann für die Eichung experimenteller Prüfmittel den Vergleichsmaßstab abzugeben. Wir besprechen in diesem Abschnitt nur die erste Aufgabe.

a) Der pädagogische Wert der Intelligenzschätzungen[1]).

Zunächst stellt die ISch. den Lehrer vor den sehr nützlichen Zwang, sich in die Individualitäten seiner Schüler zu vertiefen und das, was an psychologischen Beobachtungen im Laufe der Monate unbewußt und ungewollt angesammelt worden ist, unter einem ganz bestimmten Gesichtspunkt zu verwerten[2]). Stets empfanden die Lehrpersonen, die eine solche ISch. vornahmen, die Aufgabe als recht schwierig; zugleich aber hatten sie das Gefühl, daß sie dadurch ihre Schüler zum Teil von ganz neuen Seiten sehen und beurteilen lernten. Beides, die Schwierigkeit wie die Förderlichkeit, hat seinen Grund darin, daß die Intelligenzschätzung den Lehrer zeitweilig loslöst von der sonst ganz vorherrschenden Einschätzung der Schüler nach dem äußeren Leistungs-

[1]) Unter teilweise wörtlicher Benutzung meines Aufsatzes: (*169*).

[2]) Dies Sich-Rechenschaft-Gebenmüssen über konkrete psychologische Tatbestände ist, wie mir scheint, eine weit eindrucksvollere Einführung in die Psychologie und ihre pädagogische Bedeutung als das rein theoretische Zergliedern psychologischer Bagriffe oder das Kennenlernen psychologischer Apparate und Experimente; und ich habe daher oft Übungen zur Jugendkunde mit solchen Aufgaben der ISch. (oder ähnlichen konkreten Aufgaben der Schülerbeurteilung) eingeleitet.

Auch der Psychologieunterricht in Lehrerbildungsanstalten sollte durch solche didaktischen Hilfsmittel dafür sorgen, daß nicht nur blutlose psychologische Vorstellungen, sondern lebensvolle psychologische Einstellungen geschaffen werden.

erfolg. Jetzt soll nun einmal von der landläufigen Unterscheidung der „guten", „genügenden" und „schwachen" Schüler abgesehen werden; jetzt kann nicht mehr auf rein rechnerischem Wege aus lauter einzelnen Leistungsnummern ein Gesamtdurchschnitt und damit ein Rangplatz abgeleitet werden, sondern es soll aus unbestimmt vielen und unbestimmt flüssigen Eindrücken, die der Lehrer im Laufe der Zeit über die geistige Reife, Selbständigkeit und Anpassungsfähigkeit gewonnen hat, ein Ergebnis gezogen werden, das dem einzelnen Schüler eine bestimmte Stelle in der Reihe seiner Mitschüler anweist.

Dieser neuartigen Anforderung gegenüber werden sich verschiedene Lehrerindividualitäten sehr verschieden verhalten[1]). Der eine wird zum eigenen Erstaunen bemerken, wieviel psychologisch bemerkenswerte Erlebnisse er über seine Schüler und ihre Befähigung doch in seinem Innern aufgespeichert habe; er hatte sie sich nur nicht früher zum Bewußtsein gebracht, weil die andersartige Einstellung der Schulinteressen es nicht begünstigte; aber

---

[1]) Und deshalb hatte Binet sehr recht, wenn er hervorhob, daß die ISch, vielleicht ebense wertvolle Beiträge zur Psychologie des Lehrers wie zu der des Schülers liefere.

Binet (*4, 50*) hat einmal an zahlreiche Volksschullehrer eine Umfrage gerichtet, in der er u. a. Angaben erbat über das Verfahren, nach welchem sie zu einer Schätzung der I. ihrer Schüler zu kommen suchten. Die Antworten waren z. T. sehr lehrreich. Schon die Auffassungen des Begriffs der I. zeigten die größte Mannigfaltigkeit. Von der schulmeisterlichen Engherzigkeit, welche in der I. lediglich die Fähigkeit sieht, sich Kenntnisse zu erwerben, bis zu der feinen Erklärung einer Lehrerin: „L'intelligence ne sert pas seulement à apprendre, elle sert surtout à ‚faire sa vie'" — begegnete man allen möglichen Definitionsversuchen. Und nun die Symptome, nach welchen die Lehrer ihre Schätzung vollziehen! Da wird darauf hingewiesen, daß man der Erblichkeit nachgehen müsse, da bei Kindern intelligenter Eltern höhere I. zu erwarten sei. Empfohlen wird die Berücksichtigung des Gesichtsausdrucks; die lebhafte, offene bewegliche Physiognomie unterscheide das intelligentere Kind deutlich vom geistig trägen schwerfälligen Kinde. Einige betonen die Beobachtung beim freien Spiel; Kinder, die hier Initiative und Produktivität zeigen, seien intelligent. Der Hauptwert wird aber natürlicherweise gelegt auf das Verhalten des Kindes beim Unterricht. Hierbei wird — mit mehr oder weniger Glück — der Versuch gemacht, die eigentlich intellektuellen Momente der Unterrichtsleistungen von den mehr gedächtnismäßigen zu scheiden: Schnelligkeit der Auffassung, Fähigkeit, angewandte mathematische Aufgaben zu lösen, Verständnis für geschichtliche Beweggründe und Zusammenhänge, gute Orthographie, ausdruckvolles Lesen und manches andere wird als Symptom genannt, das den Lehrern zur ISch. dient. Endlich kommen die Lehrer auch von selbst auf die Testmethode: sie richten an die Schüler bestimmte, eigens hierfür erdachte Fragen, deren Beantwortung ihnen als Maßstab der I. dient.

es wird ihm jetzt nicht allzu schwer und ist ihm eine reizvolle mit mancherlei Überraschungen gespickte Aufgabe, aus ihnen nun Bilder der Kinderintelligenzen abzurunden. — Ein anderer wird im Gegenteil so stark an der gewohnten Einstellung haften bleiben, daß die „Güte" oder „Schlechtigkeit" der Leistungen auch für seine ISch. maßgebend bleibt; er ist nicht imstande, aus den Schulleistungen den Anteil anderer Faktoren, die nicht „Intelligenz" sind, also des Gedächtnisses, des Fleißes, des Sondertalentes auszuscheiden, sondern sieht die Leistung in ihrem komplexen Gesamtertrag als Ausdruck des kindlichen Begabungsgrades (der „Klugheit" oder „Dummheit") an. Andererseits vermag er nicht die vielen versteckten Symptome der I., die außerhalb des Schulpensums sich bekunden, gebührend zu berücksichtigen, und seine Schätzungsreihe wird letzten Endes eine nur geringfügig veränderte Neuauflage der Schulrangordnung seiner Schüler sein. — Zwischen diesen Gegensätzen gibt es nun natürlich alle möglichen Übergänge: so den Lehrer, der wohl das Bewußtsein für die Unstimmigkeit hat, die oft zwischen Schulleistungen und Veranlagungen besteht, der aber nicht die analytische Gabe besitzt, den einen Faktor von dem anderen bei seiner Beurteilung loszulösen. Und auch solche Lehrkräfte sind mir, wenn auch selten, begegnet, bei denen das Bild, das sie von den Schülern in sich tragen, von vornherein viel mehr durch deren Veranlagungsart als durch Leistungen bestimmt war, so daß sie überrascht waren, wenn der rechnerische Schulleistungsdurchschnitt zuweilen so wenig zu dem intuitiv geformten Befähigungsbild des Schülers passen wollte. — Endlich wird das Verhalten des schätzenden Lehrers natürlich auch stark durch das Fach oder die Fächer seines Unterrichts bestimmt — worauf weiter unten zurückzukommen sein wird.

Der Zwang zur Selbstbesinnung und zur Analyse seiner Berufserfahrungen ist also der erste Vorteil, den die Intelligenzschätzung dem Lehrer bringen kann.

Der zweite besteht darin, daß ihm die Schätzung zu einer neuartigen und anders orientierten Gesamtkenntnis seiner Klasse verhilft. Er erhält eine Übersicht darüber, mit welchen Verschiedenheiten der Allgemeinbegabung er in der Klasse zu rechnen hat, ob die Klasse in sich verhältnismäßig gleichartig oder ob sehr viele Abstufungen und insbesondere viele und starke Abweichungen von der normalen Mittelbreite nach oben und nach unten vorhanden sind.

Er kann drittens die Befähigungsordnung vergleichen mit anderen Gesichtspunkten, die ihm für seine Klasse zugleich zur Verfügung stehen: so mit dem Alter, der sozialen Schicht, der Schulleistung. Steht die steigende I. in direkter oder umgekehrter Korrelation zum Lebensalter? Findet sich ein Zusammenhang zwischen der Gunst bzw. Ungunst der sozialen Lage und der höheren bzw. geringeren I.? Ist die Übereinstimmung zwischen I.-Reihe und Schulleistungsrangordnung eine ziemlich große oder eine geringere? Beruht die mangelnde Übereinstimmung darauf, daß die große Mehrzahl der Schüler in beiden Reihen merklich verschieden rangiert, oder darauf, daß einzelne Schüler ganz besonders starke Widersprüche zwischen Befähigungsgrad und Schulleistung zeigen?

Die letzte Betrachtung macht ihn viertens auf einige besonders bemerkenswerte Individualitäten seiner Klasse aufmerksam, die auch in pädagogischer Hinsicht Beachtung erfordern. Denn es ist klar, daß jene — sagen wir einmal „harmonischen" — Schüler, bei denen Befähigungsgrad und Leistungserfolg im Guten oder Schlechten einigermaßen zusammenstimmen, viel leichter der unterrichtlichen und erziehlichen Behandlung zugänglich sind als jene „unharmonischen", deren allgemeine Begabung mit ihrer Leistungstüchtigkeit in Widerspruch zu stehen scheint. Da tritt denn plötzlich aus dem Schätzungsergebnis der gutbefähigte, aber wenig leistende Schüler hervor: er bedeutet eine ernste Mahnung für den Lehrer, nachzuprüfen, ob Faulheit und moralische Minderwertigkeit, oder Kränklichkeit, oder häusliche Vernachlässigung, oder Umschulung und ähnliches Ursache dieser Unstimmigkeit sei. Andererseits lösen sich die Schüler mit den typischen Willensbegabungen aus der Masse heraus: jene, die auf Grund ihres Fleißes und Pflichtbewußtseins (vielleicht aber auch auf Grund des Strebertums oder der Fähigkeit zu stumpfem, uninteressiertem Büffeln) höhere Leistungserfolge erzielen, als nach dem Maße ihrer I. zu erwarten wäre. Freilich wird die genauere Analyse auch hier zeigen, daß die Güte der Leistungen nicht so sehr in deren Selbständigkeit und Originalität als in den rezeptiven Vorzügen der Genauigkeit, Sauberkeit, gedächtnismäßigen Sicherheit beruht.

Bezog sich alles bisher Gesagte auf den einzelnen Lehrer, so treten fünftens neue Gesichtspunkte auf, wenn mehrere Lehrer die Schüler derselben Klasse schätzen. Hierzu wird ja vor allem in der höheren Schule mit ihrem Fachlehrersystem, zuweilen aber

doch auch in der Volksschule Gelegenheit sein. Natürlich muß die methodische Forderung befolgt werden, daß jeder Lehrer unabhängig vom anderen die Schätzung der Klasse vornimmt; das Ergebnis werden dann Schätzungsreihen sein, die wiederum keine völlige Übereinstimmung, zuweilen sogar ziemlich starke Unstimmigkeit zeigen. Die Unstimmigkeit hat ihren Grund einerseits in den Schätzern, andererseits in den Geschätzten. Jeder Fachlehrer ist natürlich geneigt, die I. der Schüler vornehmlich nach den Leistungen in seinem Fach bzw. in seinen Fächern zu begutachten; es wird sich also die typische Verwechslung zwischen Talent (Spezialbegabung) und I. (Allgemeinbegabung) bzw. zwischen Untalentiertheit und Unintelligenz bemerkbar machen. Dies Haften des Lehrerurteils am Fach ist zuweilen ein fast unüberwindliches. So ist mir ein Fall begegnet, daß eine Lehrkraft, die in einer höheren Mädchenschule zwei recht verschiedene Fächer unterrichtete, zunächst erklärte, nur gesondert für jedes dieser Fächer die Begabungsliste ihrer Schülerinnen aufstellen zu können; es war ihr ein ganz fremdartiger Gedanke, daß auch abgesehen hiervon so etwas wie ein allgemeiner Befähigungsgrad bei den Kindern feststellbar sein müsse. Aber gerade solche Fälle zeigen, wie wertvoll der Zwang für den Fachlehrer sein kann, sich einmal ein über die Fachleistungen hinausgreifendes Bild von den Individualitäten der Schüler zu machen; denn die Schüler sind doch einheitliche Persönlichkeiten, nicht bloße Empfänger von Prädikatnummern in Mathematik hier oder in Geschichte dort! — Unter den Schülern werden sich auch hier wieder harmonischere und unharmonischere scheiden: einerseits diejenigen, die den verschiedenen Lehrern einen ziemlich gleichartigen Eindruck bezüglich ihres allgemeinen Fähigkeitsgrades machen, andererseits die innerlich ungleichmäßigen, die auf verschiedenen Gebieten oder auch in ihrer Reaktionsweise auf verschiedene Lehrerpersönlichkeiten ganz abweichende Intelligenzbilder darzubieten scheinen; solche problematischen Naturen werden zu ganz besonderer pädagogischer Aufmerksamkeit auffordern.

Aus den Schätzungen, die mehrere Lehrer an derselben Klasse vorgenommen haben, läßt sich dann sechstens eine resultierende Intelligenzreihe der Schüler bilden, indem für jeden Schüler die Rangplätze, die er in den Schätzungen der verschiedenen Lehrer erhalten hat, zu einem Mittelwert vereinigt werden und nach diesen Mittelwerten eine neue Rangordnung aufgestellt wird.

Es ist anzunehmen, daß diese „kombinierte Reihe" das relativ zutreffendste Bild von den I.-Abstufungen der Schüler geben wird, da sie die Einseitigkeiten der einzelnen Schätzer ausgleicht.

b) Methodik der Intelligenzschätzungen.

Das vom Lehrer anzuwendende Verfahren setzt zunächst Klarheit darüber voraus, was beurteilt werden soll.

Es handelt sich nicht um ein vollständiges seelisches Gesamtbild des Schülers, sondern um eine Seite seines Wesens, die „Intelligenz", die sich nicht nur von der Willens- und Gemütsseite, sondern auch von anderen Eigenschaften seiner geistigen Beschaffenheit (wie Gedächtnis, Phantasie, Sondertalent) unterscheidet. Bei der Abgrenzung vermag die von uns vorgeschlagene Definiton der I. (S. 2/3) Dienste zu leisten.

Danach muß der schätzende Lehrer erstens achten auf die Allgemeinheit der zu beurteilenden Fähigkeit. Nicht das Benehmen des Schülers in einem Einzelfalle oder auf einem einzelnen Gebiete (z. B. der Mathematik oder der Sprache) entscheidet über die I., sondern seine Fähigkeit, in den verschiedensten Lagen und Lebensanforderungen jeweilig diejenige Denktätigkeit zu vollziehen, die der gegebenen Lage angemessen ist.

Zweitens ist die Neuheit der Anforderungen für die Einschätzung der Intelligenz maßgebend: die Fähigkeit, früher dagewesene Leistungen zu wiederholen, alte Eindrücke wieder zu beleben, ist nicht I., sondern Gedächtnis. Nun gibt es freilich im Leben keine Lage, die genau die Wiederholung einer früheren, aber auch keine, die in allen Stücken völlig neuartig wäre. Vielmehr ist Neues und Altes in den Anforderungen des Lebens stets verbunden, nur in verschiedener Verteilung. Darum hat hier der schätzende Lehrer eine analytische Aufgabe: er muß an irgend einer Verhaltungsweise des Schülers herausspüren, was und wieviel an ihr auf Gedächtnis, Übung, Gewöhnung beruht, und was und wieviel an ihr Neuorientierung, besondere Einstellung auf die besondere Konstellation der Aufgabe ist. Nur der letztere Anteil darf als Symptom der I. gelten.

Das dritte Merkmal der I., das der geistigen Anpassungsfähigkeit, schränkt sie ein gegenüber der eigentlich schöpferischen Fähigkeit, der Genialität. Diese ist ein von innen herausquellendes Streben, neue Werte zu schaffen, die der Welt ein anderes Gesicht

verleihen; I. ist dagegen doch nur ein Fertigwerden mit der Welt und ihren mannigfachen Ansprüchen. Freilich darf man, wie schon S. 5 ausgeführt, „Anpassung" nicht nur auf Gegenwartslagen beziehen, sondern muß auch das vorwegnehmende Anpassen beachten: das Planen und Abwägen, das Vorbeugen und Vorsorgen. Auch in der Jugendzeit gibt es schon genug Gelegenheiten zu vorausschauender Strategie: beim Anordnen von Spielen und beim konstruktiven Basteln, bei der planmäßigen Verteilung der häuslichen Arbeiten oder bei der Vorbereitung eines Aufsatzes, beim Ausbau von Sammlungen und — auch bei Durchführung mancher Listen und Schliche.

So muß denn der schätzende Lehrer vor allem die möglichste Vielseitigkeit der Gesichtspunkte vor Augen haben, um erst aus ihrer Gesamtheit ein angemessenes Bild der höheren oder minderen Intelligenzgrade der Schüler abzuleiten. Er wird ja in erster Linie das Verhalten der Schüler im Unterricht berücksichtigen; daneben aber wird ihm die Beobachtung der Kinder in den Pausen, auf dem Spielplatz und bei Spaziergängen wertvolle Züge liefern. Er wird vor allem auf ausgesprochene Denkleistungen achten: auf das Verstehen schwieriger Zusammenhänge, das Suchen nach Beziehungen und das Zerlegen von Zusammenhängen in ihre gedanklichen Bestandteile, das richtige Kombinieren vereinzelter Elemente, das Finden von Lösungen, das Erkennen von Fehlern und Widersinnigkeiten (Kritisieren); er wird aber andererseits auch an den Anteil der I. bei Gedächtnis- und Phantasieleistungen, beim werktätigen Schaffen, beim praktischen „Sich-zu-helfen-wissen", beim Päneschmieden usw. zu denken haben. Er wird sich vor Verwechslungen hüten müssen: daß er nicht das bloße Vorhandensein von bereitem Wissen oder eine bedeutende Sprachfertigkeit oder geistige Regsamkeit schon ohne weiteres als entsprechend hohen Grad der I. einschätze, und umgekehrt geringere Ausbildung jener anderen Eigenschaften sofort für geringe I. halte. Auch Schüchternheit ebenso wie ihr Gegenstück: der sich hochmütig in sich selbst verschließende Stolz, können oft das irrige Bild mangelnder Allgemeinbegabung vortäuschen.

Eine Erschwerung liegt ferner darin, daß lediglich Gradunterschiede der I. in der Schätzung zur Geltung kommen können; der Schätzer muß versuchen, über etwa vorhandene qualitative Verschiedenheiten der I. hinweg zu einem quantitativen Gesamturteil zu gelangen.

Man erkennt aus diesen Andeutungen die Schwierigkeit der Aufgabe, die hier gestellt ist, zugleich aber den starken Anreiz zu feineren psychologischen Analysen.

Unbedingte Voraussetzung zur fruchtbaren Durchführung der ISch. ist es, daß der Lehrer mit den Schülern mindestens schon eine Reihe von Monaten zusammen ist und Gelegenheit hatte, sie genau und vielseitig kennen zu lernen[1]). Man wird dann noch zu unterscheiden haben, ob ihn die Aufforderung zur ISch. unvorbereitet trifft, oder ob er Zeit hat, sich länger mit dem Gedanken vorbereitend zu beschäftigen. Bisher galt meist das erste: die Lehrer wurden unerwartet gebeten, zum Zweck irgend welcher psychologischen Untersuchungen baldigst eine ISch. vorzunehmen. Natürlich konnten sie dann im wesentlichen nur rekonstruierend verfahren; sie mußten sich nachträglich Beobachtungen und Erlebnisse vergegenwärtigen, die das I.-Bild der einzelnen Schüler abzustecken geeignet erscheinen. Anders liegt es, wenn sie sich selbst den Termin der Schätzung setzen; dann können sie schon in den vorangehenden Wochen die Schüler beobachten auf solche Äußerungen und Verhaltungsweisen hin, die für ihre I. kennzeichnend sind.

Das Ergebnis der Schätzung soll in einer Rangreihe der Schüler zum Ausdruck gebracht werden[2]). Um diese nach Möglichkeit exakt zu gestalten, sind verschiedene Maßregeln zu berücksichtigen.

Zunächst muß eine gewisse Auslese unter den Schülern getroffen werden:

Nur solche Schüler dürfen ihrer I. nach in ein Rangverhältnis gebracht werden, die im übrigen unter genügend gleichartigen Bedingungen stehen. Denn die feinen Unterschiede der I., welche bei der Schätzung in Betracht kommen, haben nur Bedeutung auf dem gemeinsamen Untergrunde einer sonst gleichartigen Gruppe. Deshalb hat man die ISch. fast stets auf Schüler je einer Klasse beschränkt; aber auch innerhalb der

---

[1]) Deshalb empfiehlt es sich nicht, Lehrer mit ISch. zu betrauen, die nur ein kleines Nebenfach in der Klasse unterrichten.

[2]) Natürlich ist es sehr wertvoll, wenn der schätzende Lehrer noch darüber hinausgeht und für jeden Schüler eine ausführlichere Beschreibung seiner I. entwirft. Doch fast stets wird diese Aufgabe die Arbeitskraft des Lehrers übersteigen, während die bloße Einfügung des einzelnen in eine Rangreihe leichter ohne allzu großen Zeitaufwand bewerkstelligt werden kann.

Klasse müßte man noch für die Gleichartigkeit sorgen, nicht nur indem man offenbar abnorme Schüler ausschließt, sondern indem man sich auf ein Altersgebiet bestimmter Breite beschränkt. Wenn sich z. B. in einer Klasse des fünften Schuljahres, dem normalerweise ein LA. der Schüler von 10—11 Jahren entspricht, 13jährige Knaben befinden, so sind diese nicht in die I.-Rangordnung einzubeziehen; denn der Lehrer vermag nicht festzustellen, wieviel an der bei ihnen vorhandenen I. dem höheren Alter zuzuschreiben ist; dies müßte ja beim Vergleich mit 11jährigen in Abzug gebracht werden. Deshalb siebe man jede Klasse, innerhalb deren man eine ISch. vornehmen will. Es ist nicht möglich, hierfür eine schematische Regel aufzustellen. Die Breite der Altersunterschiede, die man noch einbeziehen will, hängt von verschiedenen Umständen ab; bei höheren Jahrgängen kann man einen größeren Altersspielraum zulassen als bei jüngeren usw. Im allgemeinen ergab sich bei meinen Berechnungen, daß etwa 20—25% der Schüler jeder Klasse ausgeschaltet werden mußten.

Für die übrigbleibenden Schüler soll nun eine Rangordnung hergestellt werden. Hierbei darf aber nicht das Prinzip der Rangordnung so weit getrieben werden, daß um jeden Preis jedem Kind ein besonderer Platz zugewiesen werden müßte. Oft genug wird man es, namentlich in der mittleren Region, als eine Willkür empfinden, daß N einen schlechteren Platz als M erhalten solle, da man zu keinem eindeutigen Urteil über den verschiedenen Wert ihrer Intelligenzen gelangt. Hier gilt die Regel, daß man gleichwertigen Individuen eine gleiche Rangnummer verleiht, die dem Durchschnitt der auf sie kommenden Rangplätze entspricht. Wenn z. B. vier Individuen, welche die Rangplätze 5, 6, 7, 8 einzunehmen hätten, gleichwertig erscheinen, so erhalten sie sämtlich den Rangplatz $(5+6+7+8):4=6\frac{1}{2}$. Wenn dieses Verfahren wiederholt vorgenommen werden muß, so wird ja dadurch die Anzahl der Rangverschiedenheiten, über die man verfügt, vermindert; dieser Nachteil wird aber mehr als ausgeglichen durch den Vorzug, daß Willkürlichkeiten in der Rangierung vermieden sind. Es ist kein Unglück, wenn bei einer Klasse von 30 Schülern nur 20 oder selbst nur ein Dutzend verschiedene Rangnummern der I. vorkommen.

Notwendig ist es endlich, daß der Lehrer sich von einer etwa vorhandenen Schulrangordnung oder Fachrangordnung möglichst wenig beeinflussen läßt. Es geht nicht an, diese vorhandene Rangliste zugrunde zu legen und an ihr herumzubessern, bis man glaubt,

die I.-Reihe zu haben. Empfehlenswert ist vielmehr folgendes Verfahren: Man schreibe die Namen der zu rangierenden Schüler auf lauter einzelne Kartonblätter, die zunächst alphabetisch geordnet werden. Nun erst beginne man die Verteilung der Blätter auf verschiedene I.-Gruppen, endlich suche man innerhalb jeder Gruppe Rangplätze herzustellen. Die so gewonnene Reihe mit den dazugehörigen Nummern wird dann notiert.

Sehr wertvoll ist es, nach einem gewissen Zwischenraum (etwa 3—4 Wochen) die ganze Rangierung noch einmal von vorn zu beginnen, ohne Benutzung der zuerst gewonnenen Reihe. Der Grad der Übereinstimmung dieser beiden Schätzungen ist nach der Korrelationsmethode zu bestimmen; nur wenn der Zuverlässigkeitswert hoch ist (d. h. die beiden Reihen sehr ähnlich sind), dürfen sie weiteren Untersuchungen zugrunde gelegt werden.

Eine wirklich exakte Durchführung der ISch. verlangt von dem ausführenden Lehrer zweierlei: eine Kenntnis der psychologischen Gesichtspunkte, unter denen diese Aufgabe steht, und die Fähigkeit zu guter psychologischer Beobachtung und Beurteilung seiner Schüler. Deshalb ist es zwar erwünscht, daß sich recht viel Lehrer in ISch. versuchen, weil sie selber davon für ihre berufliche Arbeit Nutzen haben; aber dort, wo die ISch. zu weitergehenden Zwecken, insbesondere der Testeichung, verwandt werden sollen, kommen nur solche Beiträge in Betracht, die den obigen Anforderungen aufs strengste entsprechen. Wie groß die Unterschiede der Zuverlässigkeit solcher Schätzungen sind, wird im nächsten Abschnitt dargelegt.

3. Intelligenzschätzung und Schulrangordnung.

Der Vergleich der durch ISch. gewonnenen Fähigkeitsreihen mit den üblichen SchL.-Rangordnungen ergibt durchweg eine positive Korrelation von bedeutender, aber von Fall zu Fall schwankender Höhe[1]). Als Gesamtdurchschnitt kann man etwa einen Wert $\rho = 0{,}70$, also knappe „Dreiviertelsübereinstimmung“ zwischen I. und SchL. feststellen. Im einzelnen sind die Werte natürlich von den angewandten Maßwertstäben, den schätzenden Lehrern, der Art der Schule und der Besonderheit der Klassen abhängig.

[1]) Zur Methodik der Korrelationsberechnung vgl. S. 63 und das Beispiel im Anhang.

Die ersten Vergleichungen dieser Art stammen durchweg aus England; sie sind in Tabelle XVII zusammengestellt. Die SchL. wurde auf verschiedene Weise gemessen; teils benutzte man die Klassenplätze, teils den Ausfall von Schulprüfungen, wie sie in England in allen Klassen regelmäßig stattfinden. Leider werden über die spezielleren methodischen Maßnahmen bei der ISch. nicht so genaue Angaben gemacht, daß wir uns über die Sicherheit der Ergebnisse ein Urteil erlauben könnten. Nur so viel erfahren wir, daß man in einigen Fällen auf eine I.-Rangordnung verzichtete und sich mit der Herstellung von 4 oder 5 I.-Gruppen begnügte. Burt ließ zwar eine Rangordnung geben, erlaubte aber den Schätzern bedauerlicherweise, die Schulrangordnung hierzu zu benutzen und an ihr nur die etwa notwendigen Verschiebungen vorzunehmen.

Tabelle XVII.

Untersucher	Burt (*132*)	Burt (*132*)	Jones (*151*)	Gilbey & Pearson (*142*)	Waite (*173*)	Waite (*173*)	Durchschnitt aller Korr.-Werte
Schulgattung	Volksschule	höhere Schule	Volksschule	Volksschule	höhere Schule	höhere Schule	
Zahl der Schüler	30	13	249	1725	182	238	
Alter der Schüler	$12^1/_2$ bis $13^2/_2$	$12^1/_2$ bis $13^1/_2$	Korrigiert für gleiches Alter und gleiche Klasse				
Art der ISch.	Rangordnung		4 I.-Stufen	5 Intelligenz-Stufen			
Korr. der ISch. zur Schulrangordnung	—	—	—	0,67	0,74	0,63	0,68
Korr. der ISch. zu Schulprüfg.	0,81	0,78	0,70	0,68	0,82	—	0,76
Korr. der ISch. zu anderen I.-Sch. an denselb. Schülern	0,88	0,91	—	—	—	—	0,89

Die gefundenen Korrelationen sind durchweg hoch, und zwar merklich höher zwischen I. und Prüfungsausfall, als zwischen I. und Klassenplatz (0,76 gegen 0,68). Dies Ergebnis ist nicht uninteressant. Soweit es nämlich die Darstellungen ersehen lassen, wurden die ISch. vorgenommen, ohne daß der Ausfall der Schulprüfungen bekannt gewesen wäre, ja zum Teil wohl, ehe diese stattgefunden hatten. Die Schätzungen waren also nicht beeinflußt von dem Prüfungsplatz, und die ziemlich hohe Korrelation wäre als echter Ausdruck für den Grad der Übereinstimmung zwischen I. und Prüfungsleistung anzusehen.

Burt hat ferner ISch. derselben Schüler von mehreren Lehrern bzw. unbeteiligten Mitschülern vornehmen lassen; die Korr. dieser Schätzungen unter-

einander sind sehr hoch; aber da, wie erwähnt, alle Schätzer von der bekannten Schul-Rangordnung ausgingen, so ist diese Übereinstimmung nicht weiter verwunderlich und wissenschaftlich nicht brauchbar.

Die deutschen Feststellungen dieser Art sind alle aus Besprechungen hervorgegangen, die ich in meinen psychologischen Seminarübungen (früher in Breslau, später in Hamburg) dem Thema der ISch. widmete. Dies hatte den Vorteil, daß die an den Übungen teilnehmenden Lehrer aus spontanem Interesse und unter gründlichster Beachtung der in Betracht kommenden psychologischen Gesichtspunkte und methodischen Forderungen an die Arbeit gingen. Außerdem veranlaßten einige Schulleiter, die gleichfalls Mitglieder des Seminars waren, ihre Lehrerkollegien zur Vornahme von ISch. Dies führte zu einer an sich willkommenen Vermehrung des Materials, da wir nun über alle Jahrgänge von gewissen Schulen verfügten; dafür aber war die Qualität dieser Arbeiten zum Teil merklich geringer, da die Lehrer nur auf äußere Veranlassung hin die Listenaufstellung vorgenommen hatten, ohne immer das rechte Interesse und Verständnis für die neuartige Aufgabe aufzubringen. Auch wurden sie nicht genügend auf die anzuwendende Methodik vorbereitet. In einem Falle gab der Schulleiter ihnen eine schriftliche Anweisung; in den anderen Fällen erhielten sie nichts als die mündliche Aufforderung, ihre Schüler nach der I. in eine Reihe zu bringen.

Die Breslauer Schätzungen, die von den Herren Oberlehrer Scheifler und Rektor Rindfleisch (Liegnitz) veranlaßt waren, wurden von mir verarbeitet; die sehr viel zahlreicheren Hamburger Schätzungen, die sich auf Volks- und höhere Schulen beider Geschlechter erstrecken, sind von Oberlehrer Roloff (*162*) berechnet, verglichen und in ihrer Zuverlässigkeit beurteilt worden[1]).

Die Ergebnisse sind zu scheiden nach Volksschulen und höheren Schulen, da diese wegen des Fachlehrersystems unter ganz anderen Bedingungen standen als jene.

a) Die gründlichsten Schätzungen von Volksschulklassen stammen von 5 Mitgliedern des Hamburger Seminars, die unter Innehaltung aller notwendigen Maßnahmen (z. B. auch Weglassung

[1]) In den Berechnungen Roloffs sind die wahrscheinlichen Fehler nicht mitgenannt. Sie sind meist sehr gering und erreichen auch in den Höchstfällen nicht ein Drittel des Wertes ρ, so daß die Gültigkeit der Korrelation durch sie nirgends in Frage gestellt wird.

der zu alten Schüler) und mit gutem psychologischem Verständnis gearbeitet haben (Tab. XVIII). Ihre Schätzungen sind nicht einfache Ziffernlisten, sondern umfangreiche Arbeiten mit Charakteristiken der Einzelschüler, Anmerkungen über häusliche Umgebung, Hervorhebung besonderer Gründe, die auffällige Abweichungen der I.-Reihe von der Schulreihe zu erklären vermögen. Die Korrelationen schwanken bei diesen 5 Klassen zwischen 0,50 und 0,86; ihr Durchschnittswert beträgt **0,72**. Von diesen stachen die meisten anderen Schätzungen, die von psychologisch nicht besonders interessierten und vorbereiteten Lehrern im Auftrage und nach schriftlicher Anweisung des Schulleiters ausgeführt wurden, deutlich ab — auch schon in der äußeren Form, indem hier meist nur nackte Numerierungen der Kinder vorliegen. Tab. XIX enthält einen Teil dieser Schätzungen; die Korrelationen schwanken zwischen 0,59 und 0,95; ihr Durchschnitt beträgt 0,79. Aber diese Ziffer muß noch erhöht werden; denn die Lehrer hatten den methodischen Fehler gemacht, zu alte Schüler mitzuzählen. Bei der Hälfte der Klassen gelang es, diese noch nachträglich in Abzug zu bringen; es entstanden die eingeklammerten Korrelationswerte, deren Durchschnitt **0,83** beträgt.

Tab. XVIII. Zusammengestellt aus Tab. 1 und 2 von Roloff.

Volksschulklassen. Geschätzt von Mitgliedern des psychologischen Seminars.

Schuljahr . . .	1	1	2	3	5	Mittelwert
Schülerzahl . .	36	41	47	25	45	
Korr. (ρ) . . .	0,75	0,50	0,82	0,67	0,86	**0,72**

Tab. XIX.

Volksschulklassen. Geschätzt auf Grund ausführlicher schriftlicher Anweisung.

	Schule 1							Schule 2							
Schuljahr . .	4	4	5	5	6	6	Mittelwert	4	4	5	5	6	6	Mittelwert	Gesamt-
Schülerzahl	39 (29)	52	42 (38)	50	42 (29)	37		42 (32)	39	33 (29)	38	34 (29)	34		
Korr. (ρ)	0,76 (0,84)	0,73	0,95 (0,95)	0,78	0,81 (0,88)	0,74	**0,80**	0,83 (0,80)	0,59	0,74 (0,84)	0,85	0,87 (0,92)	0,81	**0,79**	**0,7 (0,8**

Eine dritte Gruppe von Schätzungen, die lediglich auf mündliche Aufforderung der Schulleiter ohne jede nähere Anweisung erfolgten, also methodisch noch tiefer stehen, stammt aus einer Liegnitzer und einer Hamburger Schule. Wir verzichten darauf, die Ergebnisse der insgesamt 25 Klassen im einzelnen aufzuführen und beschränken uns auf folgendes Hauptergebnis. Scheidet man eine einzige Liegnitzer Klasse des zweiten Schuljahres aus, die aus besonderen Gründen die ungewöhnlich niedrige Korrelation von 0,47 aufweist, so schwanken die Werte der Liegnitzer zwischen 0,78 und 0,97, die der Hamburger zwischen 0,61 und 0,94. Die Durchschnitte betragen dort 0,87, hier 0,83; aber dieser letztere Wert ist wieder durch die Hinzunahme der überalterten Schüler bedingt. Nimmt man an, daß deren Weglassung ebenso wie in der zweiten Gruppe eine Erhöhung um 0,04 herbeiführt, dann würden wir auch hier zu einem Durchschnittswert von **0,87** kommen.

Verfolgt man die 3 Durchschnittswerte 0,72—0,83—0,87, so ergibt sich mit einwandfreier Deutlichkeit, daß die Korrelationen um so höher sind, je geringer die Gründlichkeit und das psychologische Verständnis der Schätzer ist. Im einzelnen mögen besonders hohe Korrelationen auch durch andere Gründe (z. B. durch eine besondere Zusammensetzung der Klasse) veranlaßt sein, im Durchschnitt gilt zweifellos der obige Zusammenhang. Man ersieht daraus, wie stark der Lehrer im allgemeinen bei der Beurteilung seiner Schüler von den äußeren Leistungserfolgen und der darauf gegründeten Schulrangordnung beeinflußt ist; er unterliegt der Suggestion der ihm wohlbekannten Klassenrangordnung um so mehr, als er ja — im Gegensatz zum Lehrer der höheren Schule — der alleinige Schöpfer dieser Klassenrangordnung ist; denn er unterrichtet ja seine Schüler in allen oder fast allen Fächern. Derselbe Umstand also, der ihm die vielseitige Beobachtung seiner Schüler ermöglicht, erschwert ihm zugleich die Emanzipation seines Urteils von den [illegible]ein schulmäßigen Leistungsergebnissen. Es erhellt aber zugle[illegible]hieraus, wie wertvoll die Anregung zur ISch. für die psycholo[illegible] Schulung des Lehrers werden kann. Ein Lehrer, der von v[illegible] glaubt, daß die SchL. in völliger Parallele zum I.-Grad stehe, hat sich noch nie klar gemacht, welche Rolle andere Faktoren (s. S. 194) beim Zustandekommen der Schulerfolge spielen.

Besonders bemerkenswert ist in dieser Hinsicht die eine aus der Reihe fallende Liegnitzer Klasse des zweiten Schuljahres. Der Lehrer war offenbar mit großer Selbständigkeit und feinem psycho-

logischem Blick an die Arbeit gegangen, hat die Nummern in der I.-Reihe ohne Seitenblicke auf die Schulreihe gesetzt und Fälle besonderer Unstimmigkeit zwischen SchL. und I. durch besondere Bemerkungen („vom Lande zugezogen", „lange krank gewesen", „schlechte häusliche Verhältnisse") erläutert. Das Ergebnis war überraschend: eine Korrelation von nur 0,47!

Ich stellte nun an dieser Klasse noch eine andere Berechnung an, indem ich diejenigen 6 Schüler fortließ, bei welchen nach den Anmerkungen des Lehrers „besondere Umstände" vorlagen. Für die übriggebliebenen 34 stieg nunmehr die Korrelation sofort auf 0,74, näherte sich also schon sehr dem Durchschnittswerte. Daraus ergibt sich zunächst für diese Klasse, vermutlich aber auch als allgemeiner Satz: Eine etwa auftretende besonders niedrige Korrelation ist nicht dadurch bedingt, daß durchgängig die I.-Grade merklich von der Schultüchtigkeit differieren, sondern dadurch, daß für eine Minderzahl von Schülern ungewöhnlich starke Unstimmigkeiten zwischen Begabung und SchL. vorhanden sind. Diese kleine Gruppe verlangt das besondere Augenmerk des Lehrers und eine individualisierende Behandlung, denn hier ist die Gefahr am größten, daß die übliche Einschätzung der Kinder nach den SchL. zu einer falschen Beurteilung und Behandlung führt.

b) In den höheren Schulen ist die Aufgabe der ISch. für den einzelnen Lehrer bedeutend schwerer. Denn er unterrichtet nur in bestimmten Fächern und lernt daher die Schüler nur einseitig kennen; die Gefahr liegt nahe, daß er die in seinen Fächern hervortretende Sonderfähigkeit (oder -unfähigkeit) mit allgemeiner I. (oder deren Mangel) verwechselt. Da nun die Klassenrangordnung der Schüler aus der Berücksichtigung der Leistungen in allen Fächern hervorgeht, so ist es verständlich, daß die ISch. des einzelnen Fachlehrers stärker von der Schulrangordnung abweicht, als es in der Volksschule der Fall ist.

Dies wird durch Tab. XX bestätigt, die die Schätzungsergebnisse von 20 Fachlehrern an höheren Schulen Breslaus, H[illegible]gs und Bergedorfs enthält. Die Korrelation dieser Schätzungen mit dem Klassenplatz schwanken zwischen 0,33 und 0,79; ihr Durchschnittswert beträgt **0,61**, liegt also noch bedeutend unter dem tiefsten für Volksschulen errechneten Durchschnittswert 0,72.

Die Einseitigkeit der Fachlehrerschätzungen kann nun aber dadurch überwunden werden, daß wir für dieselbe Klasse

Tab. XX. Intelligenzschätzungen in höheren Schulen.

Die erste Kolumne stammt aus Breslau und ist von mir berechnet. Die übrigen Ziffern stammen aus Hamburg und Bergedorf und sind zusammengestellt aus Tab. 1 und 1a von Roloff. Die Mittelwerte der letzten Kolumne sind aus allen Werten gemeinsam gebildet.

			Gymn. Breslau	Gymn.	Ref. Gymn.	Realschule		Lyceum							Mittelwert
	Klasse . . .		U III	Quarta	Ob. III	Quinta	Quarta	VII	VI	VI	V	V	III	II	
	Schuljahr .		7.	6.	8.	5.	6.	4.	5.	5.	6.	6.	8.	9.	
	Schülerzahl		23	19	25	21	29	19	17	17	21	16	19	25	
Korrelationen	mit Klassenplatz	Lehrer A	0,43	0,70	0,37	0,56	0,45	0,57	0,75	0,70	0,63	0,58	0,65	0,70	0,61
		Lehrer B	0,55	0,78	—	—	0,46	0,79	0,75	—	—	—	0,51	0,67	
		Lehrer C	0,33	0,61	—	—	—	—	0,56	—	—	—	—	—	
		Komb. Schätzung A + B (+ C)	0,45	0,83	—	—	0,58	0,72	0,75	—	—	—	0,63	0,67	0,66
	Lehrer untereinander	A mit B	0,69	0,40	—	—	0,49	0,85	0,79	—	—	—	0,92	0,93	
		A mit C	0,65	0,58	—	—	—	—	0,82	—	—	—	—	—	
		B mit C	—	0,44	—	—	—	—	0,65	—	—	—	—	—	

Schätzungen der verschiedenen dort unterrichtenden Lehrer heranziehen und zu einer kombinierten Schätzungsreihe verbinden. Dies ist in den genannten Städten mehrfach geschehen; der eine Fachlehrer (A), der als Mitglied des psychologischen Seminars Interesse für die Aufgabe gewonnen hatte, stellte nicht nur selbst ISch. an, sondern forderte die mitunterrichtenden Kollegen[1]) (B und C) zu gleichem auf. Die aus diesen Schätzungen kombinierten Rangreihen zeigen nun eine merklich höhere Korrelation zur Klassenrangordnung, nämlich im Durchschnitt **0,66**; die Korrelationen würden wohl noch etwas höher sein, wenn nicht nur 2 oder 3, sondern alle in der Klasse unterrichtenden Lehrer zu der kombinierten Reihe beigetragen hätten.

Aber immerhin bleibt bestehen, daß in den höheren Schulen die Korrelation hinter der in den Volksschulen zurückbleibt; vermutlich beruht dies darauf, daß die schätzenden Lehrer dort nicht so sehr der Suggestion der Klassenrangordnung unterliegen. Denn keiner von ihnen ist, wie der Volksschullehrer, der alleinige Urheber dieser SchL.-Reihe; sie sind deshalb selbständiger und unbefangener in ihrer Schätzung. Durch diesen Vorzug wird der vorher genannte Nachteil der drohenden Einseitigkeit vielleicht aufgehoben.

Das Vorliegen mehrerer Lehrerschätzungen an derselben Klasse erlaubt auch einen gegenseitigen Vergleich. Wenn jeder Lehrer, wie oben angenommen wurde, einseitig durch die Erfahrung seines Faches bestimmt wird, dann muß seine Schätzung von denen der anderen Lehrer stärker abweichen, als von der (aus den Gesamtleistungen abgezogenen) Klassenrangordnung. Dies ist auch in der Tat der Fall bei der Quarta des Gymnasiums in Bergedorf. In allen anderen Klassen dagegen ist die Korrelation der Lehrerschätzungen untereinander größer als die des einzelnen Lehrers zur Klassenrangordnung. In manchen Lyzeumsklassen überstieg jene Übereinstimmung der Lehrkräfte untereinander die Höhe von 0,90! Roloff (*162*) erklärt diese Erscheinung daraus, daß die schätzenden Lehrer sich ver der Anstellung der ISch. meist mit den ebenfalls beteiligten Kollegen über verschiedene Schüler ausgesprochen haben werden, um ihre einseitigen Erfahrungen zu vervollständigen; in der Gymnasialquarta fand eine solche Verständigung nicht statt, und das Ergebnis war die geringere Korrelation. Diese Erklärung ist allerdings nicht allein ausreichend; denn bei der Breslauer Schätzung gingen die drei Lehrer in aller Unabhängigkeit voneinander vor; dennoch sind ihre Korrelationen untereinander höher als die — allerdings hier besonders niedrige — Korrelation jedes einzelnen mit der Klassenrangordnung. Die ganz hohen gegenseitigen Korrelationen im Lyzeum scheinen freilich nur durch den Gedankenaustausch der

[1]) Natürlich liegen hier ähnliche methodische Bedenken vor, wie bei den Volksschullehrern, die lediglich auf äußeren Anstoß hin die Schätzungen vornehmen.

Lehrerinnen möglich geworden zu sein. Es bedarf wohl kaum der Erwähnung, daß die Herstellung kombinierter Schätzungen nur dann einen methodischen Wert hat, wenn jede beteiligte Lehrkraft ohne vorherige Verständigung mit den anderen ihre Reihe ausarbeitet.

4. Testrangordnungen in ihren Beziehungen zu anderen Rangordnungen.

Wir kehren nun wieder zu den experimentellen Methoden der I.-Untersuchung zurück, um sie zu den bisher besprochenen Leistungs- und Schätzungsreihen in Beziehung zu setzen.

Sind wir mit Hilfe von abstufbaren Tests oder Testreihen imstande, die feineren Unterschiede der Intelligenz innerhalb einer homogenen Gruppe mit Wahrscheinlichkeit zu treffen? Die Frage ist von bedeutender Wichtigkeit; denn wird sie bejaht, so hätten wir ein Mittel, um in kürzester Zeit Urteile über den Fähigkeitsgrad von Kindern abzugeben, zu denen sonst nur eine genaue Bekanntschaft und ausgedehnte Beobachtung befähigte. Auch die Möglichkeit, Testprüfungen zur Auslese von Schülern heranzuziehen, hängt von der Bejahung dieser Frage ab.

Die Beantwortung der aufgeworfenen Frage ist aber innerhalb der experimentellen Methode allein nicht möglich. Gewiß lassen sich viele Tests finden, die sich nach theoretischen Erwägungen auf die I. beziehen und die auch fein genug abstufbar sind, um die damit geprüften Kinder in eine Stufenleiter zu bringen. Aber ob diese Stufenleiter die tatsächliche Fähigkeitsabstufung der Kinder wiedergibt, geht aus dem Experiment als solchem niemals hervor; denn es liegt in seinem Wesen, daß es gewisse rein reaktive Teilerscheinungen der I. aus ihrem Gesamtbild künstlich herauslöst (s. S. 48ff.). Wir müssen demnach die Eichungsmittel für die Tests bei einem außerhalb des Experiments gelegenen Kriterium suchen; als solches bieten sich die Urteile der Lehrer über die Fähigkeiten der Schüler an.

In den ersten Stadien unserer Arbeit zog man zu diesem Zweck einfach die üblichen Schulrangordnungen heran. Ebbinghaus (*60*) war hier im Jahre 1897 vorangegangen; viele andere folgten ihm nach. Ebbinghaus hatte an mehreren Schulen Tests verschiedener Art angewandt: Rechenuntersuchungen, Merkfähigkeitsproben, Textlückenergänzungen. Um nun festzustellen, welcher dieser Tests am besten die eigentliche „geistige Leistungsfähigkeit" wiedergebe, teilte er die geprüften Schüler nach der Schulrangordnung in die drei gleich starken Gruppen: gute, mittlere und schlechte, und

bildete für jedes Drittel die Mittelwerte der Testleistungen. Es ergab sich, daß die Ergebnisse der beiden ersten Tests von einem Rangdrittel zum anderen nur unbedeutende und regellose Abweichungen zeigten, daß dagegen die Lückenergänzung nach Menge und Güte deutlich und regelmäßig von dem guten zu dem mittleren und von diesem zum schlechten Drittel abnahm. Daraus schloß er, daß dieser Test ganz anders als die beiden anderen eigentliche I.-Unterschiede prüfe.

Dieser Schluß hat seine Berechtigung; denn wir wissen ja jetzt, daß die Schulrangordnung in bedeutender Korrelation zur I. steht. Und deshalb sind auch die Ergebnisse anderer Forscher, die sich der Einteilung der Schüler nach den SchL. bedienen, um danach den Symptomwert von Tests zu bestimmen, durchaus nicht gering zu schätzen.

Die Zahl dieser Versuche ist sehr groß; es seien hier die Namen: Meumann, Winteler, Cohn und Dieffenbacher, Binet, Stern, Minkus genannt. Sehr oft mußten die Psychologen darauf verzichten, alle Kinder einer Klasse zu prüfen, insbesondere dann, wenn Kinder verschiedener Schul- und Altersstufen untersucht wurden. In solchen Fällen ließen sie durch die Lehrer eine Auswahl treffen mit der Anweisung, gleich viele „gute" und „schwache" (zuweilen auch „mittlere") Schüler zu benennen. Daß die Lehrer bei dieser Auswahl vornehmlich die SchL. zugrunde legten, war natürlich.

Als eine der neuesten Feststellungen dieser Art sei hier noch einmal die Massenuntersuchung von Minkus (*31*) genannt, die sich ebenfalls auf die Methode der Lückenergänzung bezieht. Es waren aus mehreren Volksschul- und Fortbildungsschulklassen je acht „gute" und „schlechte" Schüler mit einem Lückontext geprüft worden, in welchem nach einem bestimmten System die Bindewörter (damit, statt dessen, weil usw.) zu ergänzen waren. Die von mir nach Minkus' Tode zu Ende geführte Berechnung erstreckt sich auf die zwei obersten Jahrgänge der Volksschule und zwei Fortbildungsschulklassen beider Geschlechter, also auf die Altersstufen 11—15. Das bereits S. 31 abgebildete Diagramm veranschaulicht die starken Unterschiede in den Testergebnissen der beiden Schülergruppen. Dies gilt vor allem in den Volksschulklassen: in der II. Volksschulklasse haben die „guten" Schüler 48% Treffer und nur $17^1/_2$% unausgefüllte Lücken, gegen $31^1/_2$% Treffer und $42^1/_2$% Lücken der „schlechten" Schüler; in der I. Volksschulklasse lauten die entsprechenden Ziffern für Treffer: 67% gegen $42^1/_2$%, für Lücken 10% gegen $21^1/_2$%. Es steht somit für die Altersstufen 11—13 Jahr der Bindewortergänzungstest in starker Korrelation zur allgemeinen Schulfähigkeit.

Die Vergleichung der Testleistung mit der SchL. hat den unleugbaren Vorzug der Bequemlichkeit; denn Schulrangordnungen

oder irgend welche ähnlichen Gradmesser für die SchL. gibt es überall; sie bieten sich als stets bereitliegende Vergleichsmittel dar. Aber ganz zufriedenstellend ist diese Vergleichsmethode doch nicht. Denn die Schulrangordnung ist nun einmal — wie der vorige Abschnitt bewiesen hat — kein eindeutiges Zeichen der Begabungsstufe; sie ist auch von Fleiß, Pflichtbewußtsein, Gedächtnis und anderen Eigenschaften abhängig; diese bringen eine Fehlerquelle in den Vergleich mit dem Testergebnis hinein.

Deshalb ist es ein zweifellos exakteres Verfahren, wenn man nicht die Schulrangordnung, sondern die Intelligenzschätzung der Lehrer an die Testrangordnung als Maßstab anlegt. Dadurch wird freilich die Untersuchungsmöglichkeit sehr eingeschränkt; denn ISch. von Lehrern sind durchaus nicht immer zu erreichen, und vor allem nicht in der Güte zu erreichen, wie sie für solche Zwecke die unumgängliche Voraussetzung bildet. Der vorige Abschnitt zeigte uns ja, daß ISch. nur dann ernsthaft zu wissenschaftlichen Zwecken verwertbar sind, wenn sie von psychologisch lebhaft interessierten und gründlich geschulten Lehrern stammen.

Liegen für eine Klasse einerseits die abgestuften Leistungen in verschiedenen Tests, andererseits die Rangreihe der ISch. vor, so lassen sich zwei Fragen mit Hilfe der Korrelationsmethode untersuchen:

1. Wie groß ist für jeden Test die Korrelation mit der geschätzten I.?

2. Läßt sich eine solche Kombination von Tests finden, daß die kombinierte Testrangordnung eine möglichst hohe Korrelation zur I. zeigt?

Die bisher vorliegenden Untersuchungen von Burt, Ries, Lobsien, Piorkowski u. a. geben nun auf beide Fragen sehr lehrreiche Antworten.

Zu 1. zeigt sich, daß verschiedene Tests sich in ihrem Symptomwert zur Feststellung der I. ganz gewaltig unterscheiden; es gibt manche, nach denen sich die Schüler ganz anders ordnen, als es die ISch. des Lehrers zeigt — solche Tests sind damit aus dem Inventar der IP. endgültig auszuscheiden. Es gibt andere, die eine bemerkenswerte Korrelation zur Lehrerschätzung aufweisen und unter diesen wieder einige wenige, bei denen die Übereinstimmung eine sehr hohe ist.

Zu 2. Jeder einzelne Test, auch ein solcher, der hohe Korrelation zur geschätzten I. zeigt, trifft doch immer nur eine be-

stimmte Seite der I. Haben wir mit mehreren Tests geprüft, so kommen diese von verschiedenen Richtungen her an die I. heran, und wegen dieser nicht vermeidbaren Einseitigkeiten bilden die Testergebnisse untereinander oft eine geringere Korrelation als jeder einzelne Test mit der ISch. Man kann nun aber diese Einseitigkeiten wiederum ausgleichen durch eine kombinierte Rangreihe, in der für jeden Prüfling die einzelnen Testrangplätze zu einem Gesamtwert vereinigt sind (s. S. 208). Ein solcher kombinierter Rangplatz entspricht besser als irgend eine Einzelleistung der — ja ebenfalls auf vielseitigen Eindrücken beruhenden — ISch.

Diese Erwägung findet sich nun in der Tat immer wieder bestätigt: durch geschickte Verknüpfung von Tests, die an sich schon eine bedeutende Korrelation zur ISch. haben, kann schließlich eine Rangreihe erzielt werden, die zuweilen eine geradezu verblüffende Übereinstimmung mit der vom Lehrer gegebenen Schätzungsreihe zeigt. Mit anderen Worten: Durch die kurzfristige Testprüfung kann unter Umständen eine Klasse bezüglich der in ihr enthaltenen Fähigkeitsgrade ebenso fein abgestuft werden, wie es der Lehrer sonst nur auf Grund einer monatelangen Beobachtung versuchen darf. Besonders optimistische Schlüsse werden hieraus von englischen und amerikanischen Untersuchern abgeleitet[1]). Ich möchte etwas zurückhaltender sein. Denn noch sind wir nicht so weit, diejenigen Tests anzugeben, deren Kombination nun ein für allemal ein mit der Lehrerschätzung stets übereinstimmendes treffsicheres Urteil gäbe. Die wenigen bisher vorliegenden Versuche sind nichts als erste Erprobungen, die durch weit umfangreichere Untersuchungen nachgeprüft werden müssen. Und noch in anderer Hinsicht muß einer gar zu optimistischen Ausdeutung dieses Satzes entgegengetreten werden: was die Tests leisten können, ist bestenfalls eine rein graduelle Unterscheidung der Kinder; die Beobachtung des Lehrers, die ja über die bloße Gradabstufung

[1]) So sagt Burt (*132*): „So sind wir mit Hilfe von einem halben Dutzend Tests imstande, eine Reihe von Knaben in eine I.-Reihe zu ordnen, welche entschieden genauer ist, als die durch Schulprüfungen zu erreichende, und wahrscheinlich genauer, als die vom Lehrer gegebene, die auf mehrjährigem persönlichen Verkehr beruht und mit ungewöhnlicher Mühe, Sorgfalt und Gewissenhaftigkeit hergestellt worden war." Und Scott (*165*) betont, daß durch ein geeignetes Testsystem ein Lehrer in den Stand gesetzt wäre, sich bei Übernahme einer neuen Klasse von der I. und dem I.-Unterschied seiner Schüler ein vorläufiges Bild zu verschaffen, das hohe Wahrscheinlichkeit habe, dem später auf Grund monatelanger Erfahrungen zu gewinnenden Bilde zu entsprechen.

hinaus auch die Fülle qualitativer Fähigkeitsunterschiede mit zu erfassen vermag, kann durch eine noch so einwandfreie Testrangordnung niemals ersetzt, höchstens ergänzt und kontrolliert werden. Aber auch dies ist schon eine wertvolle Aufgabe, die mehr als bisher in Angriff genommen werden sollte.

Wir stellen nun die wesentlichen bis jetzt vorliegenden Ergebnisse zusammen.

Als erster hat wohl Burt (*132*) in England 1909 derartige Untersuchungen veröffentlicht. Seine Prüflingszahl ist sehr gering: 30 Volks-, 13 höhere Schüler. Von den 12 angewandten Tests zeigen 6 eine Korrelation unter 0,50, 6 eine solche über 0,50 zur Schätzungsreihe. Zu den Tests mit geringerer Korrelation gehören überwiegend sensorische Tests; dies Ergebnis ist deswegen bemerkenswert, weil bei manchen Untersuchern die Neigung bestand (und auch heute noch besteht), Prüfungen der Unterschiedsempfindlichkeit zur IP. zu verwerten. Höhere Korrelationen zur geschätzten I. zeigten dagegen solche Tests, welche die Aufmerksamkeit, die Geschicklichkeit, das Ordnen und das Gedächtnis prüften; Tab. XXI gibt deren Liste. Die Kombination dieser 6 Tests aber ergibt im Vergleich mit der Schätzungsreihe die starke Korrelation von 0,85 (bei Volksschülern) und 0,91 (bei höheren Schülern) — Ziffern, die von den einzelnen Tests auch nicht annähernd erreicht worden waren.

Tab. XXI.

Burt, Versuche an normalen Kindern.

Test	Korr. der Test-Rangordnung zur ISch.	
	Volksschule	Höh. Schule
1. „Dotting" (eine aus Punkten bestehende Zickzacklinie ist während gleichmäßigen Vorbeiziehens auszupunktieren)	0,60	0,84
2. „Spot pattern" (eine Punktfigur ist nach 5maligem Zeigen im Tachistoskop nachzuzeichnen) . . .	0,76	0,75
3. „Mirror" (ein nur im Spiegel sichtbares Muster mit markierten Punkten ist auszustechen) . . .	0,67	0,54
4. „Memory" (Merkfähigkeit für konkrete und abstrakte Worte und für sinnlose Silben) . . .	0,57	0,78
5. „Alphabet" (Kärtchen mit den Buchstaben des Alphabets sind richtig zu ordnen)	0,61	0,80
6. „Sorting" (50 Spielkarten von 5 verschiedenen Farben sind in 5 Päckchen zu ordnen) . . .	0.52	0,56
Resultierende Rangordnung aus allen 6 Tests . .	**0,85**	**0,91**

In Deutschland wandte zuerst Ries (*95*) das Verfahren an. Er benutzte in 5 Volksschulklassen zwei Methoden: Merkfähigkeit für kausal verknüpfte Wörter und Assoziationen auf Grund der Beziehung von Ursache und Wirkung. Die Korrelationen zur ISch. bewegten sich (mit einer Ausnahme) zwischen 0,85 und 0,91, waren also durchweg sehr hoch. Weitere Berechnungen hatte Ries selbst nicht angestellt; aus den Rohtabellen, die er für eine Klasse veröffentlicht, vermochte ich aber noch die folgenden Korrelationswerte abzuleiten: Die beiden Tests untereinander zeigten eine geringere Korrelationshöhe (0,61); aber die aus beiden kombinierte Rangordnung besaß zur geschätzten I. die Korrelation 0,98, also fast absolute Einstimmigkeit!

In einem besonderen Buch hat Marx Lobsien (*157*) seine Untersuchungen an 40 Schülern einer Knabenmittelschule dargestellt, die er mit 11 verschiedenen Tests im Massenverfahren prüfte. Die Prüflinge waren etwa 10 Jahre alt. Das Verfahren gibt freilich zu kritischen Bedenken erheblichen Anlaß; denn die Grundsätze der Testwertung waren recht willkürliche; daher büßen die daraus abgeleiteten Rangordnungen und die zwischen ihnen gefundenen Korrelationen einen Teil ihrer Bedeutung ein.

So wurden beim Dreiwörtertest zehn Gruppen von je drei Wörtern dargeboten; die Wortgruppen waren nach der mutmaßlichen Schwierigkeit geordnet und hiernach mit verschiedenen Wertziffern versehen worden; durch Summation der Wertziffern für die gelösten Wortgruppen erhielt jeder Schüler einen Individualwert für diesen Test und damit seinen Rangplatz. — Beim Ebbinghaus-Test „Ergänzung von Textlücken" erhielten die Lücken je nach der Schwierigkeit die Wertziffern 1—5; usw.

Die Korrelationen der einzelnen Testrangreihen zur geschätzten I. waren verschieden hoch, am höchsten beim Dreiwörtertest (0,82), beim Ordnen verstreuter Wörter zum Satz (0,73), beim Ebbinghaus-Test (0,66); eine Reihe von Tests zeigte Korrelationen um 0,50 herum mit der Schätzungsreihe; am geringsten war die Korrelation (0,25) merkwürdigerweise beim Vergleich zweier konkreter Gegenstände. — Die aus allen 11 Testleistungen kombinierte Rangordnung korrelierte nun aber mit der ISch. in dem hohen Grade von 0,83. Eine Verbindung der beiden Tests mit den höchsten Einzelkorrelationen ergab die Korrelation 0,86 mit der Schätzungsreihe.

Lobsien glaubt bereits, auf Grund seiner Ergebnisse ein für 10jährige Schüler geeignetes neues Testsystem vorschlagen zu

können, das aus den oben genannten drei Tests mit den höchsten Korrelationswerten besteht. Wie mir scheint, werden aber durch die von ihm angewandte Methodik und die geringe Anzahl der geprüften Schüler nicht die Voraussetzungen erfüllt, die einen solchen Vorschlag schon rechtfertigen könnten.

Sehr hoch war die Korrelation, über die Moede und Piorkowski anläßlich ihrer Berliner Begabungsauslese berichten. Das Prüfungssystem (über dessen Zusammensetzung wir an anderer Stelle [S. 301] berichten) war außerhalb der eigentlichen Begabungsauslese lediglich zu Kontroll- und Vergleichszwecken in dem 8. Schuljahr einer Berliner Gemeindeschule durchgeführt worden (*206* S. 173ff.). Vorher hatte der Rektor eine ISch. dieser 22 Schüler vorgenommen. Die Testprüfung war recht umfangreich; die aus ihren sämtlichen Einzelwerten kombinierte Rangordnung korrelierte mit der Schätzungsreihe in der Höhe von 0,91. Der wahrscheinliche Fehler war gering[1]).

In einem gewissen Gegensatz zu dieser besonders hohen Korrelation stehen freilich die Schulerfahrungen, die mit den nach dem erwähnten Testsystem ausgelesenen Schülern in den Begabtenklassen gemacht wurden; denn bei einem Teil dieser Schüler deckte sich der Grad, in dem sie den erhöhten Anforderungen entsprachen, durchaus nicht mit dem Rangplatz, den sie in der Testprüfung erhalten hatten. Vgl. hierzu S. 303.

Auch Dr. Bobertag berichtet über Erfahrungen mit einer Testreihe, die für Zwecke der Begabungsauslese bestimmt war. Die 6 Tests (vgl. S. 282) lieferten eine kombinierte Begabungsrangordnung, für die in einer Quinta von 19 Schülern gewisse Korrelationen berechnet wurden. Die Korrelation zwischen Begabung und Klassenplatz betrug 0,71, zwischen Begabung und ISch. des Lehrers 0,91, also ebensoviel wie bei Moede und Piorkowski.

Eine besondere Stellung nimmt die Untersuchung von Susanne Engelmann (*187*a) ein, die an 19jährigen Seminaristinnen mit etwas abweichender Methodik vorgenommen worden ist. Die angewandten

---

[1]) Freilich sind bei dessen Berechnung durch Moede und Piorkowski zwei Fehler unterlaufen. Erstens ist die angewandte Formel falsch; sie setzte $wF = \pm 0{,}706 \sqrt{\frac{1-\rho^2}{2}}$, während im Nenner n stehen müßte. Denn der ganze Sinn des wF beruht darauf, daß er um so kleiner ist, je größer die Zahl der Personen ist. In der Berechnung aus dieser falschen Formel irren sich die Verfasser dann nochmals um eine Dezimale: sie schreiben 0,021 statt 0,21.

Tests schlossen sich an die Hamburger Methoden an, waren aber dem höheren Alter der Prüflinge entsprechend erschwert. Es waren die folgenden:

1. Aufsätze über einen Münchener Bilderbogen,
2. Textlückenergänzung,
3. Kritiktest,
4. Ordnungstests (Figuren verschiedener Größe, Helligkeit, Form).

Bei der Bewertung jeder Testlösung wurden auch die dabei hervortretenden psychologischen „Fähigkeiten" beurteilt. Zum Vergleich hatten nun die Lehrer der Prüflinge (denen der Prüfungsausfall unbekannt war) über dieselben „Fähigkeiten" ein Urteil abzugeben. Ergebnis: bei 12 von 35 Prüflingen war die Übereinstimmung zwischen den beiden Fähigkeitsdiagnosen „gut" (d. h. in mindestens 6 von 8 Punkten vorhanden), bei 16 „genügend" (in mindestens 4 Punkten vorhanden); somit sind in 80 % aller Fälle die Übereinstimmungen gut oder genügend. Mit Recht weist E. darauf hin, daß die Unstimmigkeiten nicht immer auf Unzulänglichkeit des Testergebnisses beruhen müßten; denn auch das Lehrerurteil war zuweilen nur unter Schwierigkeiten und Bedenklichkeiten zustande gekommen und darf daher nicht ohne weiteres als zuverlässig gelten. Psychologisch wichtig ist insbesondere der Übereinstimmungsgrad bei den einzelnen Fähigkeiten. Testergebnis und Lehrerurteil waren im Einklang: bei der Fähigkeit „selbständiges Urteil" in 77 % aller Fälle, bei „Einfühlung" in 62 %, bei „Phantasie", „Kritik", „sprachlichem Ausdruck" in 60 %, bei „Verknüpfung von Wissen und Erfahrung" in 54 %, bei „Beobachtung" und „Humor" in 51 %. — Ein Kontrollversuch an 16jährigen Schülerinnen, bei denen E. selbst auf Grund der Unterrichtserfahrungen vor Anstellung der Tests die Fähigkeiten beurteilt hatte, ergab 100 % „guter" Übereinstimmungen.

Zwei Ergebnisse seien angefügt, die mit einzelnen Prüfmethoden erzielt wurden.

Piorkowski (*90*) hat beim Dreiwörtertest (Bildung eines Satzes aus 3 Wörtern oder Wortgruppen) sehr günstige Korrelationen erzielt. Nach dem Ausfall der Prüfung mit 24 Stichwortgruppen wurden in 18 Schulklassen Rangordnungen hergestellt, die mit den gewöhnlichen Sitzordnungen verglichen wurden. Die Korrelationswerte standen durchweg (mit Ausnahme des 7. Schuljahres) über 0,50, meist um 0,70 herum; wurde an Stelle der Sitzrangordnung die ISch. durch den Lehrer gesetzt, so erreichte

die Korrelation in 15 von 18 Klassen einen noch höheren Wert. Besonders deutlich war die Beziehung des Versuchsausfalls zur geschätzten I. in den Unterklassen (3. und 4. Schuljahr).

Mit dem Definitionstest hat Roloff an 10—12jährigen Schülern noch nicht veröffentlichte Massenuntersuchungen angestellt. Wurde bei den Prüflingen auf Grund der ISch. der Lehrer ein gutes und ein schlechtes Viertel von der mittleren Hälfte abgetrennt, so ergab sich: die intelligentesten 25% hatten vor den unintelligentesten 25% auf dem Gebiet der Begriffsdefinitionen einen Leistungsvorsprung von mehr als $1\frac{1}{2}$ Intelligenzjahren — also gleichfalls eine bedeutende Korrelation.

Werfen wir schließlich einen Gesamtblick auf die Ergebnisse der Test-Schul-Vergleiche. Hierunter verstehen wir alle Korrelationsuntersuchungen, in denen Testergebnisse zu Lehrerurteilen in quantitative Beziehungen gesetzt werden — gleichgültig, ob die Lehrerurteile in Form einer gewöhnlichen Gruppierung nach SchL. oder in ISch. vorliegen. Die folgenden Angaben sind aus den Ergebnissen vieler Forscher gezogen, die nicht einzeln genannt werden können[1]).

Geringe oder schwankende Korrelation zur SchL. besitzen Tests aus den Gebieten der motorischen Leistung, der Empfindlichkeit für Sinnestäuschungen, der mechanischen Lernfähigkeit, der Rechenfähigkeit, der Suggestibilität.

Eine mittlere Korrelation — die aber von manchen Untersuchern angezweifelt wird — besitzt die sensorische Unterschiedsempfindlichkeit, die Aufmerksamkeit, die Merkfähigkeit für unzusammenhängende Einzelelemente (Ziffern, Worte usw.).

Starke Korrelation zur SchL. besitzt die Merkfähigkeit für logische Zusammenhänge, die Zuordnung („Assoziation") eines Begriffs zu einem gegebenen nach bestimmter logischer Anweisung, die Lückenergänzung, die Analogiebildung (Finden der vierten Proportionale zu drei Begriffen), die Dreiwortmethode, die Definition. Nach neuesten Untersuchungen scheinen auch Urteilstests, Kritiktests und Ordnungstests in diese Reihe zu gehören.

---

[1]) U. a. ist die Liste benutzt worden, die ich im Jahre 1911 als Anhang II der „Differentiellen Psychologie" (*12*) aufgestellt habe unter dem Titel „Test-Schulvergleichungen" (S. 486 u. 487).

Ob die Prüfmittel der praktischen I. eine bemerkenswerte Korrelation zur Schulrangordnung und zur ISch. liefern, ist noch nicht untersucht worden.

Eine Verbindung mehrerer Tests mit hoher Korrelation zu einer kombinierten Reihe liefert den Höchstgrad an Korrelation zur schulischen Reihe, der mit experimentellen Methoden erzielbar ist.

XI. Kapitel.

Die sozialen Bedingungen der Intelligenz.

Die Abhängigkeit oder Nichtabhängigkeit der I. von der sozialen Schicht stellt ein Problem dar, das nicht nur von theoretischem, sondern auch von weitgehendem praktischen Interesse ist. Denn die Einrichtung der Einheitsschule, welche eine von sozialen Bedingungen unabhängige, gemeinsame Beschulung aller Kinder vorsieht, scheint ja eng mit der Beantwortung dieser Frage verknüpft zu sein. Bedauerlicherweise hat gerade die praktische Seite des Problems bewirkt, daß an seine Bearbeitung nicht immer mit der wünschenswerten Objektivität ruhigen Forschens, sondern oft mit der Voreingenommenheit bestimmter Parteistandpunkte herangegangen wurde; und zwar ist auf beiden Seiten gesündigt worden. Indessen haben diese Erörterungen doch das Gute gehabt, mit allem Nachdruck darauf hinzuweisen, daß die Einheitsschulfrage nicht nur ihre sozialethische, sondern auch ihre psychologische Seite hat, ja daß die Einheitsschule erst dann ihre sozialethische Wirkung voll entfalten kann, wenn sie zugleich die von der Psychologie herausgestellten Unterschiede und Bedingtheiten der Begabung berücksichtigt.

Wir behandeln zuerst die ziffernmäßigen Ergebnisse, welche Prüfungen nach der BS.-Methode und andere I.-Untersuchungen bisher bezüglich der sozialen Schichtung zutage gefördert haben, sodann die möglichen Bedingungen der festgestellten Unterschiede, endlich (andeutungsweise) gewisse Folgerungen für die Schulfrage.

1. Vergleichende BS.-Prüfungen von Kindern verschiedener sozialer Schichten.

Die hierüber bisher vorliegenden Daten geben ein ziemlich buntscheckiges Bild, dem noch die volle Klarheit und Einheitlichkeit fehlt. Immerhin ist wohl eine überwiegende Tendenz in der

Richtung festzustellen, daß im Durchschnitt sozial besser gestellte Gruppen günstigere Prüfungsergebnisse lieferten als die sozial schwächeren. Kürzlich hat Karstädt (*153*) alle einschlägigen Ergebnisse in einer kritischen Zusammenstellung veröffentlicht. Seine Arbeit ist von der offen bekundeten Tendenz getragen, die gefundenen Leistungsunterschiede der sozialen Schichten als unbeträchtlich darzustellen. Man kann ihm nun darin recht geben, daß namentlich die älteren Vergleichungen — so die von Binet und Meumann — ihre Schlußfolgerungen auf methodisch unzureichendes Material gründeten. Die neueren Untersuchungen vermochte K. nicht mehr mit voller Objektivität in ihrem positiven Werte zu würdigen.

Die ersten Ergebnisse dieser Art waren dadurch gewonnen worden, daß man die Befunde verschiedener Forscher an verschiedenen Orten (ja in verschiedenen Ländern) verglich und die Abweichungen auf die verschiedene soziale Höhe der geprüften Schüler zurückführte. Dieser Schluß hat keine volle Beweiskraft; denn wenn er sich auch durchweg auf die gleiche Prüfungsmethode, nämlich das BS.sche Staffelsystem, stützte, so ist doch die Anwendung der Tests im einzelnen und die Art der Verrechnung bei verschiedenen Prüfern nie völlig gleich; außerdem können nationale, didaktische und andere Verschiedenheiten ebenfalls mitgewirkt haben, so daß es nicht angeht, die gefundenen Unterschiede lediglich der sozialen Schichtung der Kinder zuzuschreiben.

Binet war auf die Frage durch einen äußeren Anlaß gekommen. Untersuchungen von Decroly und Mlle. Degand in einer Brüsseler Privatschule (*136*) waren zu Ergebnissen gelangt, die zum Teil den Wert seiner Tests in Frage zu stellen schienen, da sich diese als durchweg zu leicht erwiesen. Von sämtlichen 45 geprüften Kindern war nämlich kein einziges unter dem I.-Stande des Alters, 9 auf diesem Stand und die übrigen über dem Niveau (13 um ein Jahr, 17 um zwei Jahr, 9 sogar um drei Jahr). Binet ist nun der Überzeugung, daß diese Zahlen keinen Beweis gegen den Wert der Tests überhaupt, sondern einen positiven Beitrag zum Studium der sozial bedingten Differenzierung darstellen. Denn die Brüsseler Kinder stammten sämtlich aus Kreisen des gebildeten Mittelstandes, die Pariser Kinder aber, auf welche die Tests zugeschnitten waren, gehörten den ärmeren Volkskreisen an. Binet berechnet hiernach den durchschnittlichen Unterschied im IA. zwischen Kindern der höheren und niederen Stände auf ungefähr

1½ Jahre. Diese Zahl kann natürlich nur einen ganz groben Annäherungswert bedeuten; sie wird vor allem bei verschiedenen Stufen des Lebensalters sehr verschieden sein, worauf Binet bedauerlicherweise keinen Bezug nimmt.

Kann man den B.schen Schlußfolgerungen eine gewisse Wahrscheinlichkeit immerhin nicht absprechen, so sind die von Meumann (*9*, Bd. II, S. 764) angeführten internationalen Vergleiche nicht mehr als ernsthafte Beweismittel anzusehen. M. zählt auf:

„1. Die Pariser Kinder waren zwar von ziemlich gemischtem Milieu, aber sie stammten im Durchschnitt aus Arbeiterkreisen der Vorstädte. 2. Am tiefsten unter ihnen stehen der I. nach die Moskauer Kinder (Frln. Dr. Schubert), die den ärmsten Arbeiterkreisen der Moskauer Vorstädte angehörten, sie stehen um volle zwei Jahre hinter den Parisern zurück. 3. Zwei Jahre über den Parisern, also vier Jahre über den Moskauer Kindern, stehen die Schüler aus den besten Petersburger Internaten (Frau Wolkowitsch). 4. Diesen ungefähr gleich — vielleicht noch etwas höher — stehen an I. die Brüsseler Kinder (Decroly und Degand), sie gehörten einem Internat mit Schülern der besten Stände Brüssels an. 5. Die von Frln. Johnston geprüften Sheffielder Kinder der Volksschule stehen denen von Binet gleich; die von ihr geprüften 25 Kinder einer höheren Schule sind den Parisern überlegen. 6. Goddards 2000 Prüflinge gehören einer im ganzen ländlichen Bevölkerung an, ihre I. ist wieder ungefähr gleich der der von Binet geprüften Kinder. 7. Genau dasselbe Bild bieten uns die italienischen Kinder: die in Rom geprüften (Jeronutti) gehören besseren Ständen an, sie stehen über den Parisern; ähnlich verhalten sich die in Turin (Treves und Saffiotti)."

Bei diesen Untersuchungen wissen wir zum Teil gar nichts über die angewandten methodischen Maßregeln; zum Teil sind die Umweltverhältnisse derart verschieden, daß die gefundenen Abweichungen gar nicht mehr zur I. Beziehung zu haben brauchen, sondern durch ganz äußerliche Gründe wie Analphabetismus usw. bedingt sein können. Die Schlußfolgerung, die Meumann aus diesem Vergleich zieht: „wir können es schon jetzt als feststehende Tatsache annehmen, daß die internationale Prüfung mit den BS.-Tests eine absolute intellektuelle Abhängigkeit des Kindes von der sozialen Lage der Eltern zeigt" — ist jedenfalls verfrüht gewesen und kann dazu beitragen, daß nun auch ernsthafte Ergebnisse psychologischer Vergleichung nicht nach Gebühr gewürdigt werden.

Wir dürfen also nur solche Befunde heranziehen, welche unter genau vergleichbaren Bedingungen und mit identischer Methodik an Kindern aus verschiedenen sozialen Schichten angestellt worden sind. Da sind nun die Möglichkeiten vorhanden, daß man

a) Kinder verschiedener sozialer Herkunft aus einer und derselben Schule, b) Kinder aus zwei Schulen derselben Schulart, aber in sozial verschieden gearteten Gegenden, c) Kinder aus zwei Schularten, deren Schüler aus verschiedenen Schichten stammen, vergleicht.

a) Bei Kindern einer und derselben Schule besteht der Vorteil, daß die Unterrichtsbedingungen durchaus gleich sind; dagegen ist es für die Lehrer sehr schwierig, die Kinder in „soziale Gruppen" zu teilen. Meist muß dann die Angabe des väterlichen Berufs ausschlaggebend sein; aber hier sind doch große Täuschungen über die damit verbundene soziale Lage möglich. Auch sind die sozialen Unterschiede innerhalb der Kinder einer Volksschule meist nicht allzu groß. Es ist daher nicht sehr verwunderlich, daß die Untersuchungen von Morlé (*159*), die auf Binets Anregung in Paris unternommen wurden, ferner die von Jaederholm (*71*, *161*) in Stockholm und die von Max Schmitt (*164*) an Zöglingen einer Würzburger Erziehungsanstalt abgehaltenen Prüfungen keine merklichen Unterschiede in den I.-Leistungen der Kinder verschiedener Gruppen zutage förderten. Etwas positiver liefen die Untersuchungen von Treves und Saffiotti (*104*, *172*) ab, denen Karstädt wohl mit Unrecht die Beweiskraft abspricht. Freilich, die Gliederung der Volksschüler in sechs soziale Gruppen geht zu weit; und es ist daher nicht zu erwarten, daß die I.-Leistung sich genau nach dieser Sechserreihe abstuft. Immerhin aber zeigt gerade die Umrechnung von Karstädt, daß bei den 6jährigen des ersten Schuljahres die drei höheren sozialen Gruppen auch die drei besten I.-Durchschnitte zeigen. Daß sich bei den 12jährigen des sechsten Schuljahres das Verhältnis zum Teil geradezu umkehrt, erklärt sich mit Selbstverständlichkeit aus den Schulverhältnissen in Rom. Nach fünf Jahren der Grundschule findet nämlich die Abwanderung an die höheren Schulen statt; es bleiben deshalb nur die ganz unbegabten Kinder der bessergestellten Kreise in der Volksschule.

b) Der Vergleich von Volksschulen gleicher Art, die aber in verschiedenen Stadtteilen gelegen sind, liefert doch schon deutlichere Ergebnisse. Hierher gehören wiederum Versuche, die Morlé auf Binets Veranlassung angestellt hat. Er verglich zwei verschiedene Volksschulen, von denen die eine im ärmsten Viertel von Paris, die andere in einer relativ wohlhabenden Gegend

lag. Aus jeder Schule wurden 30 Kinder entsprechender Altersstufen, die ohne Rücksicht auf die Schulleistungen ausgewählt waren, geprüft. Tab. XXII zeigt, um wieviel zahlreicher die rückständigen Intelligenzen in der ärmeren Schule waren. Binet berechnete hier den durchschnittlichen Vorsprung des 1A. bei den günstiger gestellten Volksschülern auf 3/4 Jahr.

Tab. XXII.

		Rückst. um 2 J.	Rückst. um 1 J.	Auf d. Niveau	Vorgeschr. um 1 J.	Vorgeschr. um 2 J.
Je 30 Volksschüler einer Schule	in sehr armer Gegend	1	11	13	4	1
	in guter Gegend . .	1	3	10	10	6

Ein stärkeres Maß des Unterschiedes tritt hervor bei Untersuchungen, welche Yerkes und Anderson in zwei Schulen von Cambridge (Mass.) angestellt haben (*178*). Sie wählten solche städtische Schulen aus, die auf Grund ihrer Lage ein sozial sehr verschiedenartiges Schülermaterial hatten; die eine bezeichneten sie als die „günstig gestellte", die andere als die „ungünstige" Schule. In jeder Schule wurden 54 Kinder beiderlei Geschlechts aus der Kindergartenklasse und dem ersten Schuljahr (meist im Alter von 5 bis 7 Jahren) geprüft, wobei möglichst für die gleichmäßige Heranziehung entsprechender Altersstufen gesorgt wurde. Der Altersdurchschnitt betrug in beiden Gruppen rund 6 Jahre. Die Anwendung der „Punkt-Skala-Methode" (s. oben, S. 156) ergab nun, daß die sozial begünstigten Knaben 37 Punkte, die anderen 29, die begünstigten Mädchen 41, die anderen 33 Punkte im Durchschnitt erhielten. Beschränkte man sich auf die am stärksten vertretenen 6jährigen, so war der Unterschied noch größer: bei den Knaben „günstig" 40 gegen „ungünstig" 29, bei den Mädchen 40 gegen 30 Punkte. Die Volksschulkinder aus einer sozial bessergestellten Gegend übertreffen also die aus dem ärmeren Stadtviertel an intellektueller Leistungsfähigkeit um 20% bis 30%.

Der festgestellte Unterschied ist ausgesprochen und deutlich — dies veranlaßt Karstädt, nach Gründen zu suchen, die seine Bedeutung herabzusetzen geeignet wären. Wie aber verfährt er hierbei? Eine Nachprüfung zeigt, daß er den englischen Text an verschiedenen Stellen mißverstanden hat und daß er infolgedessen auf einer einzigen Seite (S. 459) drei Unrichtigkeiten vorbringt, deren eine ihm sogar Anlaß gibt, mich eines unzutreffenden Zitierens zu bezichtigen.

1. K. behauptet, daß ich den Unterschied der beiden sozialen Gruppen mit der obigen Angabe „20%—30%“ zu groß darstelle: „Nach Yerkes beginnt der Unterschied bei den vierjährigen mit 20—30%, und sinkt dann bis zu 10 Jahren auf 10%.“ Wie liegt die Sache? Yerkes konnte derartiges gar nicht sagen, da bei seinem Vergleich der Kinder verschiedener sozialer Gruppen zehnjährige gar nicht vorkommen! In Wirklichkeit sagt Yerkes folgendes: Der Betrag der sozialen Differenz ist ebenso groß wie der Altersfortschritt während des 5. Lebensjahres, den er in einer anderen Untersuchung gefunden hatte; dieser Fortschritt beträgt nämlich auch 20—30%, während der Altersfortschritt bei zehnjährigen (nach dieser anderen Untersuchung) nur noch 10% beträgt. Die 10% haben also überhaupt nichts mit dem sozialen Gruppenunterschied zu tun!

2. K. sagt: „Yerkes mißt tatsächlich mehr die Unterrichtserfolge als Begabung und geistige Entwicklung.“ Als solche vom Unterricht bestimmte Tests zählt K. auf: 60—77 Wörter in drei Minuten nennen, aus drei Wörtern zwei Sätze bilden, drei Begriffsbestimmungen, rückwärts zählen.

Diese Aufgaben stehen allerdings im Yerkesschen Testsystem; aber sie werden fast nur von älteren Kindern gelöst, spielen dagegen bei 5—7jährigen, also denjenigen, die Yerkes zum Vergleich der sozialen Gruppen verwandte, nur eine ganz geringfügige Rolle. Ein Blick in die entsprechende Tabelle von Yerkes hätte genügt, um K. hierüber aufzuklären. Ich habe mich der Mühe unterzogen, die Punktzahlen, die sich auf die sämtlichen obengenannten Tests beziehen, von den Gesamtziffern abzuziehen: der Unterschied in den Leistungen beider sozialer Gruppen wird dadurch so gut wie gar nicht berührt!

3. K. behauptet, Yerkes habe es für „unfair“ erklärt, Schüler verschiedener Anstalten auf Grundlage derselben Anforderungen an die formalen Fähigkeiten zu vergleichen; er wolle gerade durch ihr Ergebnis die Unmöglichkeit solcher unfairen Vergleiche erweisen. In Wirklichkeit sagt Yerkes: „Die Feststellung der I. eines einzelnen Individuums der ungünstigen Gruppe nach Maßstäben den für die günstige Gruppe geltenden Maßstäben würde höchst unfair sein.“ Das ist durchaus zutreffend; denn das einzelne Individuum kann nur durch Unterordnung unter seinesgleichen richtig eingeschätzt werden. Unsere obigen Aufstellungen haben aber nichts mit der Einzeldiagnose, sondern nur mit statistischen Vergleichungen von durchschnittlichen Gruppenwerten zu tun; die Problemstellung ist also eine durchaus andere und hat überhaupt keine Beziehung zu dem, was Yerkes unfair nennt.

Die kurze Mitteilung zum gleichen Thema, die J. und R. Weintrob (*175*) über vergleichende Testuntersuchungen an mehreren Neuyorker Schulen machen, ist ohne wissenschaftliche Bedeutung. Die Verfasser wollen zwischen Schulen, die sich aus ganz verschiedenen sozialen Schichten rekrutierten, keinen intellektuellen Höhenunterschied gefunden haben. Nun ist zunächst, wie schon Bobertag bemerkte, die Berechnungsmethode durchaus willkürlich; vor allem aber sind die geprüften Schulen nicht nur sozial, sondern auch rassenmäßig geschieden. Die Schule aus der sozialen

Unterschicht war zugleich eine rein jüdische, die Schülerschaft der anderen stammte größtenteils aus der rein amerikanischen Mittelschicht. Hier kann also leicht die intellektuelle Frühreife der jüdischen Rasse den sozial bedingten Niveauunterschied ausgeglichen haben.

Über die Bedeutung der Umwelt für das IA. liegt eine Würzburger Untersuchung von M. Schmitt (*164*) vor, der ich gleichfalls keinerlei Beweiskraft beizumessen vermag. Er prüfte 100 Kinder beiderlei Geschlechts im Alter von 4—16 Jahren, die, aus niederen Kreisen stammend, Insassen von Internaten waren, nach der BS.schen Anordnung von 1911. Hat die Umwelt — so argumentiert Schmitt — einen Einfluß auf das IA., so müssen die Kinder der Erziehungsanstalt bei IP. besser abschneiden als die Kinder der unteren Volksschichten, die in ungünstigerer, nämlich der gewöhnlichen häuslichen Umwelt weilen. Zu diesem Zweck wären natürlich vergleichende Untersuchungen von Kindern beider Gruppen nötig gewesen. Sch. verzichtet aber hierauf, nimmt vielmehr als feststehend an, daß sich die in normaler (häuslicher) Umwelt lebenden Kinder nach ihrer I. symmetrisch verteilen; und da nun unter den Anstaltszöglingen sich mehr Kinder mit I.-Rückstand als mit Vorsprüngen befanden, so habe jedenfalls die günstigere Umwelt keinen Einfluß ausgeübt. An diesem Beweisgang sind schon die Voraussetzungen falsch. Erstens trifft es nicht zu, daß die Binet-Simon-Methode in der alten Form symmetrische Verteilung der Intelligenzen ergab (s. o., S. 158ff.); wir wissen durchaus nicht a priori, ob normale Würzburger Schulkinder, nach gleicher Methode wie die Anstaltskinder geprüft, nicht in noch größerer Häufigkeit I.-Rückstände als jene gezeigt hätten. Sodann aber befanden sich unter den Anstaltskindern Prüflinge bis zu 16 Jahren; es ist aber bekannt, daß für diese höheren Jahrgänge das BS.-System nicht mehr ausreicht und daher zu tiefe I.-Werte ergibt. Hiergegen bietet auch eine Veränderung der Berechnung, die Schmitt anwendet, nicht genügenden Schutz. Vielleicht also beruhen die vielen I.-Rückstände, die Schmitt bei den Anstaltskindern fand, auf diesen zu alten Prüflingen; keinesfalls können sie als Vergleichsmaterialien für die Frage des Umwelteinflusses verwendet werden.

c) Vergleich verschiedener Schulgattungen. Der Unterschied der sozialen Lage ist am stärksten ausgeprägt bei den Schülern der Volksschulen einerseits, der höheren Schulen anderer-

seits; und Vergleichungen dieser beiden Gruppen versprechen daher die wichtigsten Ergebnisse. Um so merkwürdiger ist es, daß hier bis vor kurzem erst ein einziger Versuch vorlag, bei dem man von wirklicher Vergleichbarkeit reden kann[1]).

Im Jahre 1912 hat eine Gruppe Breslauer Lehrer unter meiner Beratung eine Untersuchung veranstaltet zur systematischen Vergleichung von Volksschülern mit den aus gehobenen Kreisen stammenden Insassen einer Vorschule[2]). Es sollte festgestellt werden, ob zwischen gleichaltrigen Kindern beider Gruppen typische I.-Unterschiede bestehen und welche Größe diese auf verschiedenen Altersstufen haben. (In Preußen berechtigte damals dreijähriger Besuch der Vorschule, aber erst vierjähriger der Volksschule, zum Eintritt in die Sexta des Gymnasiums; auch das sollte geprüft werden, inwieweit diese Maßregel nicht nur durch den hier und dort bewältigten Lehrstoff, sondern auch durch die allgemeine geistige Reife der Kinder psychologisch gerechtfertigt war.)

Es wurden fünf Gruppen geprüft, die bezüglich des Alters sorgfältig vergleichbar gemacht wurden, 7- und 9jährige Vorschüler, 7-, 9- und 10jährige Volksschüler, im ganzen ungefähr 150 Knaben. Um eine Verständigung der Prüflinge untereinander zu verhindern, wurde folgendermaßen verfahren: Vier Experimentatoren (mit ihren Protokollführern), die alle auf die Technik in gleichmäßiger Weise eingeübt waren, prüften an demselben Nachmittag in verschiedenen Räumen. Jeder Prüfer konnte in dieser Zeit 4—5 Prüfungen erledigen, und jeder Prüfling mußte nach seiner Untersuchung sofort nach Hause gehen; so wurden 16 bis 20 Klassengenossen geprüft, ohne daß ein Gedankenaustausch zwischen ihnen möglich gewesen wäre.

In Tab. XXIII sind die Gesamtwerte der IA. vereint und zugleich von mir in IQ. umgesetzt.

Dem IA. nach sind sowohl die 7jährigen wie die 9jährigen Volksschüler durchschnittlich um ein halbes I.-Jahr im Rückstand

[1]) Die schon weit zurückliegenden Untersuchungen von Miß Johnston (*150*) in Sheffield, die neben 193 Volksschülerinnen auch 25 Schülerinnen eines Lyzeums prüfte, sind methodisch so unzureichend, daß sie weder im positiven Sinn, wie es Meumann tat, noch im negativen, wie es Karstädt versucht, für unsere Frage verwertet werden können. Wir erfahren nichts darüber, ob die beiden Gruppen in bezug auf Alter überhaupt vergleichbar waren; wir hören, daß bei der Prüfung Binets Pariser Bilder und Texte ohne die für englische Verhältnisse nötige Umgestaltung verwandt wurden usw.

[2]) Vgl. hierzu: Hoffmann *146, 147*; W. Stern *167, 168*.

hinter den Vorschülern gleichen Alters; die 10jährigen Volksschüler haben ungefähr das gleiche IA. wie die um ein Lebensjahr jüngeren Vorschüler.

Tab. XXIII.

Vergleich von Kindern verschiedener sozialer Schichten.
(Nach A. Hoffmann [*146*, *147*].)

	7jähr.		9jähr.		10jähr.	Alle	
	Volksschüler	Vorschüler	Volksschüler	Vorschüler	Volksschüler	Volksschüler	Vorschüler
Anzahl	37	33	25	32	29	91	65
Durchschnittliches IA.	8	8,5	10,3	10,8	10,9		
Durchschnittlicher IQ.	1,14	1,21	1,13	1,20	1,09	1,12	1,20

Das Maß des IQ. bewährt sich auch hier; liefert es doch typische, vom Alter fast unabhängige Werte für die beiden sozialen Schichten. Die drei Altersgruppen der Volksschüler haben IQ. um 1,1, die beiden Vorschulgruppen um 1,2. Innerhalb jeder Altersgruppe beträgt der Rückstand der Volksschüler 0,07, d. h. $^1/_{13}$ der Normalintelligenz[1]).

In der öffentlichen Erörterung dieser Untersuchung ist öfters ihr Hauptergebnis schlagwortartig in dem Satz zusammengefaßt worden: 10jährige Volksschüler standen auf der geistigen Höhe von 9jährigen Vorschülern. Diese Formulierung ist, wie Karstädt mit Recht hervorhebt, nicht glücklich; denn die 10jährigen Volksschüler der Breslauer Prüfung scheinen durch irgend einen Zufall eine etwas schwächere Gruppe gebildet zu haben. Das geht hervor aus dem etwas niedrigen IQ., sowie daraus, daß sie die um ein Jahr jüngeren Kinder ihrer eigenen Schulform nicht um ein ganzes, sondern nur um 0,6 I.-Jahre übertrafen. Es ist nun aber durchaus ungerechtfertigt, aus diesem Umstand mit Karstädt eine Ablehnung der ganzen Untersuchung ableiten zu wollen. Denn nicht der Vergleich der 10jährigen mit den 9jährigen ist das wesent-

[1]) Daß die Durchschnitts-IQ. sämtlicher fünf Gruppen mehr als 1 betragen, liegt wohl vornehmlich an der Auswahl der Schüler; es waren nämlich nur solche Schüler geprüft worden, deren Alter mit dem Klassenalter übereinstimmte; Sitzenbleiber, verspätet Eingeschulte und Umgeschulte waren fortgelassen worden. Außerdem mögen auch in der benutzten Staffelordnung nach Bobertag die Testansätze für die jüngeren Kinder noch etwas zu leicht sein. Dadurch wird aber das Ergebnis der Vergleichung in keiner Weise berührt; uns kommt es ja nur auf den Abstand beider Gruppen an.

liche, sondern der Vergleich gleichaltriger Gruppen beider Schulformen untereinander; und der hier gefundene Vorsprung von einem halben I.-Jahr, der sowohl bei den 7jährigen wie bei den 9jährigen Prüflingen hervortrat, ist von der Prüfung der 10jährigen gänzlich unabhängig.

Ein weiterer Einwand Karstädts geht dahin, daß auch Kinder einfachster Kreise die höhere Schule besuchen, und zwar gerade die begabtesten unter diesen, so daß der Vorsprung der Vorschüler vielleicht von diesen, nicht von den Kindern der gehobenen Schichten stamme. Nun war aber damals noch der Prozentsatz der höheren Schüler aus einfachen Kreisen zu gering, um Gruppenwerte stark beeinflussen zu können; vor allem aber wurden solche Kinder fast stets erst von der Volksschule aus in die Sexta eingeschult, sind daher in der Vorschule so gut wie gar nicht zu finden. Auch dieser Einwand ist demnach nicht stichhaltig.

Um festzustellen, welche psychischen Fähigkeiten es im speziellen sind, in denen die Leistungen beider Gruppen weit voneinander abweichen, stellte ich aus den Tabellen Hoffmanns alle diejenigen Proben zusammen, in welchen die Häufigkeit richtiger Lösungen bei den Vorschülern mindestens um die Hälfte die Häufigkeit bei den gleichaltrigen Volksschülern übertraf. Die so gewonnene Liste (Tab. XXIV) ist nach dem steigenden Wert dieses Häufigkeitsverhältnisses geordnet.

Der eingeklammerte Test ist nicht zu berücksichtigen, weil der Ausfall durch schulische Bedingungen herbeigeführt ist[1]); im übrigen kann nur bei dem Aufsagen der Monate eine unmittelbare Unterrichtswirkung mitgespielt haben. Alle sonstigen Tests betreffen entweder den Umfang der Merkfähigkeit oder eigentliche Denkleistungen, insbesondere solche, die auch eine gewisse Beherrschung der Sprache voraussetzen: Erklärung von Begriffen und Bildern, Sätze ordnen und ergänzen, Verstandesfragen, Kritik von Widersinnigkeiten. (Aber auch der stumme Denktest „5 Gewichte ordnen" ist dabei.) Da zudem die Merkfähigkeitsproben zweifellos nicht sowohl das mechanische Gedächtnis prüfen, sondern die Fähigkeit der Erfassung und Verknüpfung einer Inhaltsreihe mit der Aufmerksamkeit und (bei den Erinnerungen an Gelesenes und den 26silbigen Sätzen) das Sinnverständnis, so können wir sagen, daß es fast durchweg formale Leistungen der eigentlichen Denkfähigkeit sind, in welchen die Kinder gehobener

[1]) Im ersten Schuljahr der Volksschule wird der Zahlenkreis nur bis 20 durchgenommen; die Aufgabe überschreitet diese Grenze.

Stände die gleichaltrigen Kinder des Volkes übertreffen. Es ist wohl kein Zufall, daß fast alle Tests des BS.-Systems, die als besonders wertvoll gelten, in der Liste stehen[1]), die minderwertigen dagegen fehlen.

Tabelle XXIV.

Tests mit den größten Leistungsunterschieden bei Kindern verschiedener sozialer Schichten.

Test	Häufigkeit richtiger Lösungen bei 7jährigen Volkssch.	Vorsch.	Bei den Vorschülern wievielmal so häufig?
5 Gewichte ordnen	31½ %	50 %	1,6
6 Erinnerungen an Gelesenes wiedergeben	37 „	56 „	1,6
3 leichte Verstandesfragen	51½ „	87½ „	1,7
6 Ziffern nachsprechen	30 „	62 „	2,1
Bildertest, Stufe der Erklärung	8 „	18,2 „	2,3
Begriffserklärung über Zweckangaben hinaus	18 „	56 „	3,1
(80 Pf. auf 1 Mark herausgeben	16½ „	59½ „	3,6)
	bei 9jährigen		
3 Reime finden	54 %	81 %	1,5
6 Ziffern nachsprechen	56 „	84 „	1,5
26-silb. Sätze nachsprechen	48 „	75 „	1,6
Monate aufsagen	58 „	96 „	1,7
Lücken eines Textes ergänzen	32 „	59 „	1,8
60 Worte in 3 Minuten nennen	20 „	40 „	2,0
Abstrakte Begriffe erklären	25 „	62½ „	2,5
Kritik absurder Sätze	28 „	75 „	2,7
Durcheinandergemischte Worte zu einem Satz ordnen	16 „	48 „	3,0

Kennzeichnend für die Tendenz Karstädts ist das Verfahren, durch welches er versucht, das Ergebnis dieser Spezialvergleichung möglichst zu entwerten. Vereinzelte Testlösungen oder Versager, die in dem Hoffmannschen Bericht verstreut unter vielen anderen Proben auftauchen, werden herausgesucht und so aneinander gereiht, daß der unbefangene Leser glauben muß, sie seien typische Musterbeispiele der ganzen Prüfung gewesen. Durch diese Beispiele soll dann bewiesen werden, daß der Vorsprung der Vorschüler auf ihrer besonderen grammatischen Übung, der Nachteil der Volksschüler auf sprachliche Mißverständnis, Schüchternheit und ähnlichem beruhe. Selbstverständlich wird kein besonnener Beurteiler bestreiten, daß diese Faktoren mitspielen. Auch ich selbst habe schon in der ersten Besprechung der Ergebnisse ausdrücklich darauf hingewiesen, daß die Verschiedenheit der äußeren Bedingungen stark mitwirkt

[1]) Ein vorläufig nicht verständliches Paradoxon bildet der Test der „schweren Verstandesfragen"; in ihm zeigen die 9jährigen Volksschüler mit einer Häufigkeit von 32 % sogar eine kleine Überlegenheit gegenüber den nur 25 % der Vorschüler.

(*167*, *168*, S. 122). Aber es ist nicht gerechtfertigt, durch eine tendenziöse Auswahl und Zusammenstellung von Beispielen beim Leser ein durchaus irriges Bild von der Bedeutung dieser Beispiele zu erwecken[1]).

2. Weitere Intelligenzvergleichungen von Kindern verschiedener sozialer Schichten.

Wir haben nun noch über einige weitere Ergebnisse sozialer I.-Vergleichungen zu berichten, die nicht mit der Staffelmethode, sondern mit Einzeltests oder auf nichtexperimentellem Wege erzielt worden sind.

Bei drei aus meinem Seminar hervorgegangenen Untersuchungen war die Wirkung der sozialen Differenzierung zwar nicht das Hauptproblem, wohl aber ein wichtiges Teilproblem, auf das bei der Auswahl der Prüflinge von vornherein Rücksicht genommen wurde. Drei ganz verschiedene Altersstufen wurden bei diesen Arbeiten berücksichtigt.

An kleinen Kindern im Alter von 4—8 Jahren haben G. und A. Schober (*108*) Untersuchungen mit der Heilbronnerschen Serienmethode (s. S. 95) angestellt. Die Umrißzeichnung eines Gegenstandes ist auf einer Reihe von Karten in ständig zunehmender Vollständigkeit wiedergegeben; die Karten werden hintereinander vorgelegt, und es wird geprüft, wann die Kinder den vorgelegten Gegenstand erkennen, ob sie die geringen Unterschiede von Karte zu Karte bemerken und benennen, und ob sie zum Schluß anzugeben imstande sind, was wohl noch weiter hinzugefügt werden könnte. Die aus Warteschulen, Kindergärten und verschiedenen Schulen stammenden 60 Kinder gehörten zur Hälfte sozial günstigeren Schichten an; von jeder Altersstufe stammten je drei Knaben und Mädchen aus besser gestellten, je drei aus sozial weniger günstigen Kreisen. Diese Ziffern sind zu gering, um endgültige Schlüsse zu erlauben; immerhin ist die Gesamtrichtung der Ergeb-

---

[1]) Dazu kommen psychologische Fehldeutungen wie die folgende: „Die Ordnung von Wörtern zu einem Satze ist für Vorschüler mit grammatischer Schulung nur Übungsanwendung, da sie sofort an der Endung, Gegenstand, Ergänzung und Beifügung erkennen (ein, verteidigt, Herrn, mutig, Hund, treuer, seinen)." Die schwere geistige Leistung, den Zusammenhang des Satzes durch Umordnung der Wörter herzustellen, ist etwas so durchaus anderes als die Übung im Beachten und Verwenden von Endungen, daß keinerlei Recht besteht, eine so unbewiesene Behauptung dogmatisch als psychologische Selbstverständlichkeit hinzustellen.

nisse unverkennbar; während bei den jüngeren Jahrgängen (4 bis 6 Jahre) ein fast durchgängiger, zum Teil recht bedeutender Vorsprung der sozial gehobenen Kinder besteht, findet vom 6. zum 8. Jahre eine Annäherung beider Gruppen statt, die zuweilen sogar zu einem kleinen Vorsprung der Kinder aus einfacheren Kreisen führt. Bedingt ist dieser Wechsel durch einen besonders starken Altersfortschritt der sozial Schwachen in den ersten Schuljahren. Für diese ermöglicht also erst die Schule eine Entwicklung, die bei den anderen unter günstigeren häuslichen Bedingungen Aufgewachsenen schon früher eingesetzt hatte[1]). Der Rückstand der jüngeren Kinder aus sozial ungünstigeren Verhältnissen ist nicht etwa vorwiegend durch sprachliche Unbeholfenheit bedingt, denn bei der Benennung der Bildunterschiede ist die Gruppendifferenz unbedeutend; es sind vielmehr die Funktionen des Erkennens und der phantasiemäßigen Ergänzung, in denen sie erst später zur Entwicklung kommen.

Eine (noch nicht veröffentlichte) Massenuntersuchung von H. P. Roloff hat besonders genau die sozialen Unterschiede der Prüflinge berücksichtigt. Angewandt wurde der Definitionstest mit 14 Begriffen an 9—13jährigen Gymnasiasten, Realschülern und Volksschülern aus Hamburgs Nachbarstadt Bergedorf. Es ergaben sich zwischen den Leistungen der drei Schülergruppen sehr bemerkenswerte Unterschiede; und zwar waren diese zwischen Gymnasiasten und Realschülern bedeutend größer als zwischen diesen und den Volksschülern. Rechnete man nämlich die Leistungen in I.-Jahre um, so betrug der Durchschnittsvorsprung in der Definitionsleistung:

zwischen Gymnasiasten und Volksschülern 3 Jahr,
zwischen Gymnasiasten und Realschülern 2¼ Jahr,
zwischen Realschülern und Volksschülern ¾ Jahr.

Es ist nun unwahrscheinlich, daß die beiden höheren Schulen in den unterrichtlichen Bedingungen so stark voneinander abweichen, daß die Minderleistung der Realschüler daraus erklärt werden könnte. Der Grund ist vielmehr in der Verschiedenheit der Schichten zu sehen, aus denen die Schüler herstammen. Die Be-

[1]) Natürlich ist diese Annäherung beider Gruppen im 7. und 8. Jahre nur bezüglich derjenigen geistigen Funktionen auszusagen, die durch den Test geprüft werden. Der Test bietet den älteren Kindern keine bedeutende Schwierigkeit mehr und läßt etwa sonst vorhandene I-Unterschiede nicht mehr hervortreten.

rufe der Väter sind nämlich der Mehrheit nach folgendermaßen zu kennzeichnen:

bei den Gymnasiasten: Großkaufleute, Akademiker, höhere Beamte,
bei den Realschülern: Handwerker und Gewerbetreibende, niedere Beamte und Angestellte,
bei den Volksschülern: gelernte und ungelernte Arbeiter.

Auch eine Statistik der Wohnweise wurde angestellt; sie ergab bemerkenswerterweise einen größeren Abstand zwischen Gymnasiasten und Realschülern, als zwischen diesen und den Volksschülern. Es wohnten:

	Gymnasiasten	Realschüler	Volksschüler
in Einfamilienhäusern und Villen	76%	25%	—
in Etagenhäusern	24%	75%	100%

Mit viel älteren Jahrgängen (12—16jährigen) hat es die Untersuchung zu tun, die Minkus an Breslauer Volks- und Fortbildungsschulen mit mehreren Tests angestellt hat (*31*). Er hatte für jedes Geschlecht zwei sozial verschiedene, im übrigen parallele Schulen ausgewählt: eine Volksschule aus besonders armer Gegend wurde einer anderen aus günstigerer Lage gegenübergestellt; die gewerbliche Fortbildungsschule trat der kaufmännischen gegenüber, die, wie man annahm, vornehmlich von Schülern und Schülerinnen aus besserer sozialer Lage aufgesucht wurde. Bei der Bearbeitung der gewonnenen Materialien wurde diese Gegenüberstellung dadurch verstärkt, daß aus jeder Schule noch eine Auslese solcher Schüler getroffen wurde, bei denen der Stand des Vaters einen besonders deutlichen Schluß auf günstige bzw. ungünstige Gestaltung der Umwelt zuzulassen schien. So entstanden Vergleichsgruppen von je 10 Prüflingen für jede Klassenstufe jeden Geschlechts.

Bisher liegen ziffernmäßige Vergleichsergebnisse nur für den einen Test der Bindewortergänzung aus 4 Klassenstufen vor, worüber bereits an anderer Stelle dieses Buches (S. 31) kurz berichtet wurde. Für unser Problem zeigt dort die rechtsstehende Figur, daß in den zwei oberen Volksschulklassen die Leistungen beider sozialer Gruppen durchaus gleichwertig waren, daß dagegen in der Fortbildungsschule ein Rückstand der Kinder aus ungünstiger Umwelt auftritt, der von der Unterstufe zur Mittelstufe noch zunimmt. Tab. XXV gibt die für beide Geschlechter vereinigten Werte der prozentualen Häufigkeit der richtigen Ausfüllungen.

Tabelle XXV.

Trefferprozente	Volksschulklasse		Fortbildungsschule	
	II.	I.	Unterstufe	Mittelstufe
Günstige Umwelt . . .	37	52	58½	67½
Ungünstige Umwelt . .	37½	51	49	51

Dieser Ausfall der sozialen Vergleichung ist um so bemerkenswerter, als er zu dem Ausfall einer anderen Vergleichung in einer gewissen Gegensätzlichkeit steht. Die nach Schulleistungen gruppierten Schüler zeigten nämlich in ihren Testleistungen bedeutende Unterschiede, und zwar war der Rückstand der schlechten (unbegabten) Schüler gegenüber den guten in den Volksschulklassen noch größer als in der Fortbildungsschule.

Für diese Befunde habe ich folgende Erklärung zu geben versucht (*31*, S. 46):

„Bei den 12—14jährigen, denen also ihrer Entwicklung nach die Aufgabe der Bindewortergänzung größere Schwierigkeiten macht, ist die intellektuelle Fähigkeit für den Ausfall entscheidend, nicht solche Eigenschaften, die stark von der Umwelt beeinflußt werden. Bei den höheren Jahrgängen dagegen, die dem eigentlich logischen Teil der Aufgabe, der Erfassung der Zusammenhänge, besser gewachsen sind, tritt nun ein Faktor hervor, der stark durch die äußeren Verhältnisse mitbestimmt ist: das ist die sprachliche Ausdrucksfähigkeit. Hierzu mag noch ein Gewöhnungsfaktor kommen. In der Schule sind die Kinder, gleichgültig aus welcher sozialen Schicht sie stammen, durch den Unterricht im weiten Umfang an sprachliche (mündliche und schriftliche) Äußerung ihrer Gedanken gewöhnt. Späterhin verliert sich diese Gewöhnung, und zwar um so mehr, je weniger der gewählte Beruf zu feiner differenzierten Sprech- und Schreibweisen veranlaßt. Daher haben die Besucher der kaufmännischen Fortbildungsschule (Kontoristinnen und Textilwarenhändler) mehr sprachliche Übung als die der gewerblichen Fortbildungsschule (Schneiderinnen und Gürtler)."

Zum Schluß seien die interessanten Untersuchungen des früheren Bremer Schulinspektors Hartnacke (*144*a u. b) erwähnt, die ohne Zuhilfenahme experimenteller Methoden mit Statistik und Lehrerurteil arbeiten. In Bremen gab es damals (1917) neben den unentgeltlichen Volksschulen auch „entgeltliche", die ein mäßiges Schulgeld nahmen; Pensum und Methodik war durchaus gleich; nur die soziale Herkunft der Schüler war infolge des Schulgeldes eine andere. H. weist nun nach, daß die Schülerschaft der sozial günstigeren entgeltlichen Schulen auch intellektuell höher stand.

Zunächst ergab die Versetzungsstatistik der Jahre 1913—1916, daß in den unentgeltlichen Schulen 8%, in den entgeltlichen nur 3% das jeweilige Klassenziel nicht erreichten und deshalb sitzen-

blieben. Daß der Grund hierzu nicht allein in Verwahrlosung und häuslicher Vernachlässigung der ärmeren Schüler, sondern zum großen Teil in wirklichen Fähigkeitsunterschieden zu suchen sei, geht aus einer Erhebung mit anderer Fragestellung hervor. Im Herbst 1916 wurden die Lehrer der 6. Klassen (3. Schuljahr) aufgefordert anzugeben, welche Schüler sie ihrer Denkfähigkeit nach für fähig halten würden, in eine höhere Schule überzugehen[1]); es waren 96 von den 836 Schülern der entgeltlichen, 41 von den 1413 Schülern der unentgeltlichen Schulen, also dort 11,5%, hier nur 2,9%. Eine dritte Statistik bezieht sich auf die 50 Kinder beider Schulen, welche nach dem Lehrerurteil die allerhöchste Stufe der Denkfähigkeit (Prädikat I) besitzen: die Väter dieser Kinder hatten zur Hälfte „geistige" Berufe, sie waren Beamte, Kaufleute, Handlungsgehilfen, Bauführer usw.; unter der anderen „werktätigen" Hälfte ist der an absoluter Zahl so starke Arbeiterstand nur ganz spärlich vertreten.

Fassen wir zusammen. So buntscheckig das Bild zunächst erscheint, so lassen sich doch einige allgemeinere Züge herausheben. Wurden Kinder verschiedener Schulen, deren Zöglinge durchschnittlich aus sozial und materiell verschieden gestellten Schichten stammten, mit identischen I.-Methoden geprüft oder beurteilt, so ergab sich in den meisten — und zwar gerade in den wissenschaftlich am besten vorbereiteten — Fällen ein durchschnittlicher Vorsprung der sozial bevorzugten Gruppe[2]). Das Er-

[1]) Es handelte sich nicht etwa um die tatsächliche Anmeldung; diese ist von der Zustimmung der Eltern abhängig, die namentlich bei den Schülern der unentgeltlichen Schulen nur selten gegeben wurde. Von den oben genannten 41 intelligenten Schülern der unentgeltlichen Schulen wurden nur 4 für die höhere Schule angemeldet.

[2]) Karstädt gibt eine andere Zusammenfassung, die ein durchaus irreführendes Bild liefert. Er rechnet nämlich mechanisch jede Untersuchung ohne Rücksicht auf ihren wissenschaftlichen Wert und auf die der Vergleichung zugrunde gelegten Gruppen = 1 und kommt zu folgender Zählung: „Von 20 Untersuchungen scheiden 4 aus; von den übrigen 16 zeigen 6 eine positive Beziehung zwischen Begabung und Gesellschaftsklassen, 4 keine und 6 eine negative Beziehung; d. h. nach 6 Ergebnissen waren tieferstehende soziale Schichten den höherstehenden an Begabungsleistungen überlegen." Welche 6 Untersuchungen diese letzte Gruppe bilden, wird nicht klar; mir ist keine wissenschaftlich ernst zu nehmende Untersuchung bekannt, die als Gesamtergebnis einen Vorsprung der sozial ungünstigen Gruppe gezeigt hätte. (Daß einzelne Teilergebnisse gelegentlich einen solchen Vorsprung zeigen, soll natürlich nicht geleugnet werden.)

gebnis bezieht sich mit großer Wahrscheinlichkeit vor allem auf die jüngeren Altersstufen (etwa bis zum 12. Jahre); für die höheren Jahrgänge liegen erst dürftigere und in sich weniger gleichartige Materialien vor, so daß man hier noch nicht von einer Endgültigkeit des Ergebnisses sprechen darf.

Der wahrscheinlich gemachte Unterschied besteht aber lediglich massenstatistisch als Durchschnittswert der verglichenen Gruppen; innerhalb jeder Gruppe ist die Streuung ziemlich groß, so daß stets einige Kinder der sozial ungünstigen Gruppe über den Durchschnitt der sozial günstigeren Gruppe herausragen und umgekehrt einige Kinder der günstigen noch unter dem Durchschnitt der ungünstigen stehen.

3. Die Ursachen der Unterschiede.

Nun aber erhebt sich erst die Frage nach den Ursachen für den Vorsprung der sozial gehobenen Gruppen, und damit stehen wir vor einem der schwierigsten, heute noch nicht endgültig zu lösenden Probleme der I.-Forschung.

Wenn die IP. tatsächlich von der nackten I.-Disposition als solcher allein abhängig wäre, dann läge ja ein eindeutiges Ergebnis vor: es wären dann die Kinder der gehobenen Stände im Durchschnitt mit einer stärkeren I.-Anlage begabt als die Kinder der Masse. Nun wird es aber allseitig zugestanden, daß die IP. trotz aller dahingehenden methodischen Bestrebungen doch nicht die reine Anlage feststellt, sondern diese in Verbindung mit all den Einflüssen, unter denen der Prüfling bis zum Moment der Prüfung gestanden hat; und deshalb ist auch meines Wissens von keiner Seite behauptet worden, daß der gefundene Unterschied einfach und ausschließlich als angeborener Begabungsunterschied der beiden Schichten aufgefaßt werden müsse. Freilich wäre es genau ebenso unwissenschaftlich, wollte man aus einer vorgefaßten Meinung heraus die Möglichkeit einer solchen innerlich bedingten Verschiedenheit von vornherein leugnen und alles nur auf äußere Faktoren, insbesondere solche der Schulbildung zurückführen. Das Wahrscheinlichste ist, daß beide Faktoren, der innere und der äußere, zugleich an dem Ergebnis beteiligt sind — ohne daß wir den verhältnismäßigen Anteil eines jeden schon im einzelnen klar übersehen können.

Am äußeren Faktor müssen wir aber wiederum zwei Momente unterscheiden: Schule und Elternhaus. Karstädt und manche andere sind geneigt, lediglich die Verschiedenheit des Schulbetriebes und Unterrichts für die geringere Prüfungsleistung der sozial ungünstiger gestellten Gruppe verantwortlich zu machen. Nun ist aber der Gruppenunterschied nicht nur zwischen Volksschulen und Vorschulen hervorgetreten, sondern auch zwischen Schulen gleicher Gattung, die nur durch die verschieden günstige Lage oder durch Schulgelderhebung voneinander abwichen. Es ist sehr unwahrscheinlich, daß sich diese Schulen so stark im Lehrplan usw. unterschieden, um daraus allein den Abstand der Prüfungsleistungen erklärlich zu machen — zumal ja die Anforderungen der Prüfungen nur wenig mit unmittelbarem Schulwissen zu tun hatten. Und eben der letzte Grund verbietet, daß wir die in Breslau festgestellte Verschiedenheit gleichaltriger Vorschüler und Volksschüler allein auf den abweichenden Unterrichtsbetrieb („Schulsystem, Lehr- und Stundenpläne, Besuchsziffern", Karstädt [*152*, S. 839]) schieben dürften. Die durchschnittlichen Klassenfrequenzen betrugen bei unseren Breslauer Untersuchungen in den parallelen Klassen der Volksschule 58, der Vorschule 48 — der Unterschied ist nicht besonders groß. Die Abhängigkeit vom Lehrplan kann nicht bedeutend sein; denn gerade die in Tab. XXIV zusammengestellten Tests, welche die deutlichsten Leistungsunterschiede bewirkten, haben (mit einer Ausnahme, siehe Anmerkung 1, S. 235) nichts mit Schulkenntnissen und Gedächtnisstoffen, sondern mit formalen Denkfähigkeiten zu tun. Es würde schließlich nur der Unterschied der Unterrichtsmethode übrigbleiben, die nach Karstädt in der Vorschule größeren Wert auf sprachliche Formulierung der Denkinhalte legt und dadurch die Lösung der Denktests erleichtert hat.

Aber gerade in der letztgenannten Hinsicht ist die Schule nicht als selbständiger Einfluß, sondern nur als Teil des kindlichen Gesamt-„Milieus" zu werten. Die Fülle geistiger Anregungen, welche auf Kinder der gehobenen Stände vom ersten Lebensjahre an einströmt — durch Gespräche und Bilderbücher, Spiele und Rätsel, Reisen und Vergnügungen —, muß ihre intellektuelle Leistungsfähigkeit bedeutend fördern, auch solchen Kindern gegenüber, welche die gleiche intellektuelle Anlage haben, aber entsprechender Anregungen entbehren. Das geht mit besonderer Deutlichkeit aus den S. 237 beschriebenen Schoberschen Untersuchungen mit vor-

schulpflichtigen Kindern hervor. Die geistige Atmosphäre der sozialen Umgebung, insbesondere des Elternhauses, ist daher in stärkerem Maße für den geistigen Höhenunterschied verantwortlich zu machen als der verschiedenartige Schulunterricht; ja die Verschiedenheit des Schulunterrichts ist zu einem großen Teil selbst erst dadurch bedingt, daß durchschnittlich Kinder aus ganz verschiedenen geistigen Umwelten in die Volksschule und die Vorschule kommen.

Obige Auffassung erfährt eine bemerkenswerte Bestätigung durch eine eingehende, aus Lehrerkreisen stammende, Individualanalyse eines geprüften Kindes (*154*). O. Kosog hat seinen fünfjährigen Sohn nach BS. geprüft und kam zu dem Ergebnis, daß ihm nach den Leistungen ein IA. von 9 Jahren zugeschrieben werden müßte; seine wirkliche Begabung entspricht aber durchaus nicht diesem abnormen Wert. Kosog ist nun auf Grund der genauen Kenntnis des Knaben in der Lage, bei vielen Tests die besonders guten Leistungen plausibel zu machen, und zwar wesentlich aus Einflüssen des Elternhauses. Die Unterstützung, die das Zahleninteresse des Knaben bei den Eltern fand, mannigfache Rätsel- und Reimspiele, die mit ihm getrieben wurden usw., bewirkten, daß die entsprechenden Testaufgaben für ihn keine neuen Anforderungen bedeuteten; aber auch bei den eigentlichen Denkoperationen machte sich zweifellos bemerkbar, daß die geistig sehr rege Umgebung eine gewisse formale Denkübung bei ihm begünstigt hatte.

Soviel ist freilich zuzugeben, daß die Tests dem Einfluß des gebildeten Elternhauses bzw. der Unterrichtsmethode der Vorschule in höherem Maße, als es wünschenswert ist, entgegenkommen. Eine Nachprüfung der Ergebnisse mit stummen Tests, welche die natürliche „praktische" Intelligenz ohne Zuhilfenahme sprachlicher Formulierungen feststellen, wäre daher eine dringliche Aufgabe; vielleicht, daß sich dann das Bild etwas ändern würde[1]).

Bei aller Würdigung der genannten äußeren Umstände, welche die Leistung der sozial schlechter gestellten Kinder herabzusetzen geeignet wäre, darf man aber doch nicht übersehen, daß auch innere Gründe, d. h. solche der wirklichen Veranlagung, mitspielen können[2]). Denn die soziologischen Tatsachen sind mit biologischen eng verknüpft; die soziale Schichtung hängt mit Vererbungsmomenten zusammen. Zu einem großen Teil sind die Kin-

---

[1]) Freilich lieferte der im Staffelsystem enthaltene, stumme Test „5 Gewichte ordnen", einen starken Ausschlag zugunsten der Vorschüler.

[2]) Hartnacke hat versucht, die Bremer Lehrerschaft zu Schätzungen der Gründe anzuregen, die zum Sitzenbleiben der Kinder führten. Es ergab sich, daß weit häufiger innere Momente (Schwächen der Willens- oder Denkveranlagung) als äußere Momente genannt wurden (*144c*).

der der sozialen Oberschicht die Nachkommen von Generationen, die sich eben durch besondere Tüchtigkeit und Befähigung aus den unteren Schichten heraufgearbeitet haben; sie sind also bereits das Erblichkeitsergebnis einer schärferen Auslese, die wenigstens teilweise intellektueller Natur ist. Das Grundgesetz aller sozialen Bewegung, daß die oberen Schichten nach oben hin absterben und sich ständig aus den Tüchtigkeitsreserven von unten her auffüllen, hat augenscheinlich auch hier gewirkt[1]). Natürlich bedeutet das nicht, daß nun jedes Kind der begünstigten Schichten die Anwartschaft auf höhere I. mit in die Wiege bekomme, denn zum Teil sind ja die Triebkräfte, die den Aufstieg ihrer Vorfahren bedingten, ganz anderer als intellektueller Natur; und zu einem anderen Teil hat die Tüchtigkeit der Vorfahren durch Überalterung und Entartung der Familien längst ihre erbliche Wirkungskraft verloren. Nur so viel wird man auf Grund der Erblichkeitsbedingungen annehmen dürfen, daß unter den Kindern der gehobenen Schichten stärkere Intelligenzen in relativ größerer Häufigkeit vorkommen werden als unter denen der Masse. Andererseits ergibt sich gerade aus dieser biologischen Betrachtung ein Gesichtspunkt, der auch mit der praktischen Erfahrung und den Ergebnissen der IP. übereinstimmt: Unter den Kindern der breiten Volksschichten befinden sich stets, wenn auch vielleicht in geringerer Häufigkeit, Individuen von besonders hoher Fähigkeit, eben jene, die zum sozialen Aufstieg berufen sind, die Eltern der künftigen führenden Schichten.

Neben der absoluten Begabungshöhe kann auch das verschiedene Entwicklungstempo die Abweichung in den Leistungen der sozialen Gruppen bedingen[2]). Bei Kindern der gehobenen Schichten geht zuweilen die geistige Entwicklung mit einer größeren Schnelligkeit vor sich als bei den Kindern des Volkes, ohne daß sie darum in ihrem Endergebnis zu einem höheren Begabungs-

[1]) Schallmeyer, „Vererbung und Auslese", sagt: „Unablässig, wenn auch sachte, werden die oberen Stände durch überdurchschnittlich tüchtige Individuen und Geschlechter aus den unteren Ständen ergänzt. Wer ohne Vorurteil an diese Frage herantritt, wird kaum zu dem Glauben kommen können, daß in dem jahrhundertelangen wirtschaftlichen Wettkampf der Familien Begabungsunterschiede gar keinen Einfluß auf deren soziales Emporkommen, Untenbleiben oder Herabgleiten geübt haben." Einer exakten Messung entziehen sich ja die geistigen Erbanlagen, weil sie stets nur im Verein mit anderen Faktoren ... in Wirkung traten." (S. 293, zitiert nach Hartnacke.)

[2]) Hierauf hat besonders Meumann (*9*, S. 766) hingewiesen.

niveau führen müßten; so mag der experimentell gefundene Vorsprung zum Teil nur das Kennzeichen einer gewissen Frühreife sein, die noch nicht zu eindeutiger Prognose verwertet werden kann.

4. Pädagogische Ausblicke.

Da die Aufgabe dieses Buches eine psychologische ist, so können wir hier auf die pädagogische Seite des Problems „Intelligenz und soziale Schichtung" nicht näher eingehen. Die unparteiische und objektive Behandlung der psychologischen Tatsachen und Ursachen hatte sich ja gerade hier als besonders dringlich erwiesen, weil von anderer Seite durch zu frühe Verquickung der Betrachtung mit tendenziösen Wünschen die reine Tatsachenfeststellung erschwert worden war. Aber zum Schluß dieses Kapitels und abseits von der wissenschaftlichen Berichterstattung sind wohl einige Andeutungen über die Folgerungen gestattet, die aus dem Vorangegangenen für die künftige Gestaltung der Einheitsschule zu ziehen sind.

Der Unterbau der Einheitsschule soll alle Kinder ohne jede Unterscheidung vereinen. Ohne Unterscheidung nach Stand und Geldbeutel der Eltern, aber auch ohne Unterscheidung nach Fähigkeiten. Da gerade in diesem frühen Alter die Kinder der sozial gehobenen Schichten oft einen geistigen Vorsprung vor ihren Altersgenossen aus dem Volk haben, so wird der gemeinsame Unterricht für viele der ersten Gruppe eine gewisse Minderung des Nutzwertes bedeuten, der sonst in gleicher Zeit für sie zu erzielen wäre. Aber dieser Verlust an Unterrichtstempo, Förderung der geistigen Entwicklung und Menge des aufzunehmenden Lehrgutes soll mehr als aufgewogen werden durch die sozialethische Bereicherung, die das Gemeinschaftsleben mit Kindern aus allen Volksschichten bedeutet. Freilich: die Voraussetzung dafür, daß man jene intellektuelle Verzögerung nicht schwer zu nehmen braucht, ist eine doppelte: erstens darf diese Unterrichtsgemeinschaft ohne Gliederung nach Fähigkeiten nicht zu lange währen; zweitens muß bei eintretender Differenzierung dafür gesorgt werden, daß die vorhandenen Fähigkeiten nun voll gewürdigt und in die ihnen gemäße Schulbahn gelenkt werden.

Bezüglich der Dauer der Grundschule sind ja die Ansichten noch sehr geteilt. Mir scheint ein sechs Jahr währendes Zusammenbleiben aller Kinder eine schwer zu rechtfertigende Beeinträch-

tigung der begabteren Kinder zu sein; denn je stärker mit steigendem Alter sich die Befähigungsgrade und -arten differenzieren, um so mehr wirkt der noch andauernde gemeinschaftliche Unterricht als eine Hemmung und Verschleppung für diejenigen, die einem anderen Unterrichtstempo gewachsen wären. Diese so beeinträchtigten Kinder würden zu einem merklichen Prozentsatz gerade die Kinder der gehobenen Stände sein; Entsprechendes gilt aber natürlich für die Hochbegabten des Volkes. Mir will es daher angemessen scheinen, daß spätestens nach vier Jahren die erste Gliederung der Kinder nach Begabungen erfolge, der dann später weitere Differenzierungen zu folgen hätten.

Über den Anteil, den die Kinder der gehobenen sozialen Schichten an diesen oberen Verzweigungen der Einheitsschule (den Nachfolgern der heutigen „höheren" Schulen") zu nehmen hätten, sei es erlaubt, eine anderweitig veröffentlichte Äußerung (*170*, S. 11) hier nochmals wiederzugeben:

„Das genealogische Alter an sich, also der bloße Umstand, daß eine Familie bisher schon zu den oberen Schichten gehört hatte, gibt keine Anrechte mehr auf besondere Beschulung und Berufszulassung. Für die öffentliche Schulpolitik dürfen nur noch die individuellen Fähigkeiten des Geistes und Charakters und die auf Grund dieser Fähigkeiten zu erwartende künftige Tüchtigkeit zu wertvollen Leistungen bestimmend sein. Alte Familienkultur kann hieran Anteil haben, sofern sie sich als Fähigkeit und Tüchtigkeit in einzelnen Nachkommen bekundet. In doppelter Weise kann die Zugehörigkeit zu einer solchen Familie die Befähigung des Kindes beeinflussen: von innen her durch Vererbung wertvoller geistiger Eigenschaften — eben jener Eigenschaften, die früher der Familie den Aufstieg zur sozialen Höhe und das Verbleiben auf ihr ermöglicht hatten —; von außen her durch die besondere geistige Atmosphäre des Elternhauses, die das Kind vom ersten Lebensaugenblick an umgibt. Sind diese Bedingungen im Kinde lebendig, so wird es den Wettbewerb mit den begabten Kindern aus der Masse des Volkes gut bestehen und mit diesen hinaufsteigen zu den Höhen der Bildung und der Berufsstellung; und ich zweifle keinen Augenblick daran, daß bei dieser rechtmäßigen, durch und durch demokratischen Konkurrenz der verhältnismäßige Anteil der Kinder aus bisher kulturtragenden Familien ein beträchtlicher bleiben wird. Es sind dann eben diejenigen „alten" Familien, die noch „jugendlich" genug sind, um die Altersschätze geistigen Erbgutes in junge

frische Gegenwartskraft umzusetzen. Und es wird ein Glück sein für den Kulturfortschritt, daß dieses lebenskräftige Geisteserbgut, welches die Kontinuität mit der Vergangenheit wahrt, das unhistorische Neuerungsstreben der neu in die Kultur Eintretenden durchsetzt und abmildert. Ja noch mehr; auch bereichert wird der Kulturfortschritt durch dieses Nebeneinanderwirken von fähigen Individuen verschiedener Schichten; denn die Tüchtigkeit der Kinder älterer Familien wird zum Teil auch ein qualitativ anderes Gepräge tragen als die der neu Emporsteigenden; Züge, die diesen fehlen, werden dort mehr ausgeprägt sein und umgekehrt. Die künftige Auslese der Tüchtigen wird eine Beweglichkeit, Feinfühligkeit und Vielseitigkeit besitzen müssen, die alle Arten der Fähigkeit zu ihrem Rechte kommen läßt, nicht etwa nur die, die dieser oder jener sozialen Schicht besonders eigen sind. Hier wird die Psychologie der Begabungsauslese zugleich zu einer sozialpolitischen Angelegenheit von höchster Wichtigkeit werden.

Jene alten Familien freilich, bei denen das Alter sich nicht mehr in gegenwärtige Tüchtigkeit umsetzt, gehen des Anspruchs verlustig, eine besondere Rolle im öffentlichen Leben zu spielen; sie sind überaltert und müssen der genealogischen Verjüngung Platz machen."

Allerdings muß jene „Verjüngung" auch unterschiedlich verstanden und gewertet werden. Denn gerade heute finden wir unter den sozial gehobenen Schichten neben den Familien mit älterer Bildungs- und Kulturtradition die sich immer breiter machenden Vertreter des „neuen Reichtums", die zu einem großen Teil sehr unerfreulichen Eigenschaften des Intellekts und des Charakters ihren plötzlichen Wohlstand verdanken. Wenn die Kinder gerade dieser Kreise in besonders großer Anzahl auf Grund des elterlichen Reichtums in die führenden Berufe und Stellungen gelangen sollten, so könnte dies zur größten Gefahr für die kulturelle und sittliche Höhe der nächsten Generation werden. Doppelt dringlich wird gerade gegenüber den Kindern dieser Schichten die Forderung einer entschiedenen, von der Allgemeinheit zu regelnden, auf Geist und Charakter zugleich gerichtete Tüchtigkeitsauslese sowie die Loslösung der Schulbahn vom Geldbeutel der Eltern.

IV. Teil.

Psychologische Methoden der Schülerauslese.

XII. Kapitel.

Methodische Grundsätze für die Mitwirkung der Psychologie an der Schülerauslese.

An keiner Stelle ist bisher die Psychologie in eine so unmittelbare und verantwortliche Beziehung zu praktisch-pädagogischen Aufgaben des öffentlichen Lebens getreten, wie bei der Auslese der Befähigten. Um so mehr ist es nötig, daß sie sich über die anzuwendende Methodik strenge Rechenschaft gibt und alle Verzweigungen der vielseitigen Aufgabe, vor die sie gestellt ist, durchdenkt. Es ist dies nicht immer mit genügender Sorgfalt geschehen, was aus der Plötzlichkeit verständlich wird, mit der die neuen Forderungen befriedigt werden mußten. Wohl haben wir Psychologen diese Forderungen kommen sehen[1]); aber wir haben auch darauf hingewiesen, daß die Psychologie, um ihnen ganz gewachsen zu sein, über Institute, Arbeitsmittel und Arbeitskräfte von ganz anderem Umfange als bisher verfügen müßte. Nun wurde gleichsam über Nacht die gegliederte Einheitsschule, die Förderung der Begabten, die Auslese der Geeigneten für Schulen und Berufe zur Wirklichkeit, und die Psychologie mußte, ohne daß jene Voraussetzungen erfüllt waren, mithelfen; denn Begabung und Eignung sind psychologische Kategorien; und ihre Feststellung ist eine mindestens zum Teil psychologische Tätigkeit. Erst an der praktischen Arbeit selbst konnten sich die neuen Methoden ausbilden, wobei Mißgriffe und Einseitigkeiten zunächst nicht ganz zu vermeiden waren; und auch jetzt sind wir noch mitten in den ersten Stadien dieser weitausschauenden Entwicklung.

Nicht das ganze Problem gehört in den Rahmen der gegen-

---

[1]) Meine bereits 1915 geschriebene Programmschrift „Jugendkunde als Kulturforderung" (Leipzig, Quelle und Meyer) beschäftigt sich eingehend mit jenem Problem. Programmatisches enthält ferner mein 1917 gehaltener Vortrag (*234*, *235*).

wärtigen Betrachtung. Die rein pädagogischen, sozialen und wirtschaftlichen Seiten der Frage gehen uns hier nichts an, sondern nur die Hilfsdienste, welche die Psychologie mit ihren Methoden zu leisten hat. Ferner sind die Eignungsprüfungen an Erwachsenen und alle diejenigen Methoden, die sich allein auf die Feststellung einer Sonderfähigkeit (z. B. der technischen Begabung) beziehen, auszuschalten. Unser Thema ist: Welche Rolle spielen psychologische Methoden bei der Feststellung einer allgemeinen Hochbefähigung von Kindern und Jugendlichen, die für besondere Schulbahnen oder Berufsvorbereitungen ausgelesen werden sollen?

Die Feststellung der Intelligenz kommt bei der Schülerauslese in doppelter Weise in Betracht. Einmal kann es sich darum handeln, daß die geistige Allgemeinbegabung das entscheidende Kriterium für die Auslese darstellt; in anderen Fällen sollen zwar die Kinder und Jugendlichen auf Grund von Sonderbegabungen ausgelesen werden, aber daneben muß auch ihre allgemeine Intelligenz mitberücksichtigt werden. Dort steht demnach die Intelligenzuntersuchung im Mittelpunkt, hier ist sie nur ein sekundäres Mittel der Auslese.

Im allgemeinen wird bei jüngeren Kindern mehr der erste, bei älteren Jahrgängen der zweite Gesichtspunkt vorwalten. Denn bei Kindern im Alter von 9—11 Jahren wird es noch kaum möglich sein, schon eine qualitative Begabungs- und Interessenrichtung mit einiger Sicherheit und vor allem mit prognostischer Wahrscheinlichkeit feststellen zu können; wohl aber ist es hier schon möglich, über den Gesamtgrad der geistigen Regsamkeit, über die allgemeine Höhe der geistigen Anpassungsfähigkeit ein Urteil zu fällen. Je älter die Kinder sind, um so mehr tritt zu dem graduellen Unterschied auch der qualitative; typische Begabungsrichtungen, die zu bestimmten Schulgattungen und Berufsvorbildungen zu prädestinieren scheinen, machen sich immer deutlicher bemerkbar, und die früher ziemlich farblose Allgemeinintelligenz stellt sich mehr und mehr in den Dienst besonderer Interessengebiete. Hiernach werden auch die psychologischen Intelligenzmethoden einzurichten sein: bei den jüngeren wird man mehr die Gesamthöhe, bei den älteren mehr die besondere Färbung der Intelligenz („sprachliche Intelligenz", „technische Intelligenz" usw.) zu beachten haben.

Bedeutende Schwierigkeiten macht die Frage, wie sich die Psychologie mit ihrer Mitarbeit in die Gesamtheit der bei der

Auslese beteiligten Faktoren einordnet. Als solche Faktoren sind zu nennen: 1. Die öffentliche Instanz (Staat, Gemeinde), die für eine möglichst richtige Verteilung der Kinder auf die Schulen zu sorgen hat. „Möglichst richtig" bedeutet hier aber: einerseits den Fähigkeiten und Neigungen der Schüler angemessen, andererseits den bestehenden Möglichkeiten der Einschulung und den künftig zu erwartenden Gliederungen der Berufe angemessen. 2. Die abgebende Schule, aus der die auszulesenden Kinder stammen. 3. Die annehmende Schule, in die sie hinein sollen. 4. Die Eltern, die mit Recht ein gewichtiges Wort bei der Bestimmung des Schulschicksals mitsprechen wollen; denn das Schulschicksal ist die wichtigste Vorbedingung des künftigen Berufs- und Lebensschicksals. 5. Die Schüler selbst, die, insbesondere wenn es sich um höhere Altersstufen handelt, nicht als bloße Objekte öffentlicher Maßnahmen, sondern als Willenssubjekte mit eigenen Bestrebungen bewertet werden müssen. 6. Die Psychologie.

Das Verhältnis der ersten fünf Gruppen untereinander kommt für unsere gegenwärtige Erörterung nicht in Betracht; in ihrer Gesamtheit stehen sie als „pädagogische" Faktoren dem letzten „psychologischen" Faktor gegenüber. Und da sind bei den bisherigen Veranstaltungen der Schülerauslese bald die rein pädagogischen Gesichtspunkte als allein maßgebend betrachtet worden, so von allen denen, die mit den bisher üblichen Methoden der Schulzeugnisse und Aufnahmeprüfungen auszukommen suchen; bald ist mit starker Einseitigkeit der psychologische Gesichtspunkt in den Vordergrund geschoben worden — so vor allem von Moede und Piorkowski in Berlin und ihren Nachfolgern; endlich wird ein organisches Ineinanderwirken der pädagogischen und psychologischen Hilfsmittel bei den verschiedenen Hamburger Auslesen angestrebt. Es wird in folgendem zu zeigen sein, daß weder ein rein pädagogisches, noch ein rein psychologisches Verfahren die Aufgabe befriedigend zu lösen vermag. Wir brauchen ein durchgebildetes „pädagogisch-psychologisches Verfahren", dessen Programm sich folgendermaßen formulieren läßt:

> Die Verantwortung, die mit der Auslese verknüpft ist, ist so groß, daß alle verfügbaren Hilfsmittel in der denkbar vollkommensten Weise verwendet werden müssen. Die Vertrautheit mit der Kindesseele, die Lehrer und Eltern besitzen, und der Schatz an Beobachtungen, der von ihnen in langem Umgang mit dem Kinde angesammelt worden ist, müssen bei der Voraus-

lese zur Geltung kommen; daneben müssen Prüfungen den gegenwärtigen Stand der Fähigkeiten und Leistungen feststellen. Bei beiden Methoden, der Beobachtung wie der Prüfung, muß der Pädagoge mit dem Psychologen zusammengehen. Die Entscheidung muß auf einer Gesamtberücksichtigung aller gewonnenen Kriterien beruhen; sie muß insbesondere in den weiten Grenzgebieten der zweifelhaften Fälle feinfühlig individualisierend vorgehen und sich nicht auf mechanische Ziffernwerte verlassen; sie muß ebenfalls in gemeinsamer Arbeit von Pädagogen und Psychologen erfolgen. Auch späterhin muß fortlaufend eine Erprobung der Auslesemaßnahmen stattfinden; es muß festgestellt werden, ob die ausgelesenen Schüler den psychologischen Prognosen entsprechen und ob die angewandten Methoden sich bewährt haben.

Die Psychologie muß beanspruchen, zu dieser Arbeit als unentbehrliche Mitarbeiterin herangezogen zu werden; sie darf aber andererseits nicht vergessen, daß sie hier nur Hilfsdisziplin für eine dem Wesen nach pädagogische und soziale — nicht schlechthin psychologische — Kulturaufgabe ist. Es muß ferner eine wirkliche Gemeinschaft beider Gebiete angestrebt werden. Darin ist wahrlich nicht das Ideal zu sehen, daß schulfremde psychologische Wissenschaftler von außen her in die Schule eingreifen und von den Lehrern mit Mißtrauen als Eindringlinge betrachtet werden. Heute ist dies noch nicht ganz zu vermeiden, weil psychologische Einstellung und Schulung in der Lehrerschaft noch wenig verbreitet ist. Es ist aber zu hoffen, daß die Pädagogen bald die Notwendigkeit ihrer Mitwirkung auch an dem psychologischen Teil der Auslesearbeit einsehen werden, und daß sie mit steigender psychologischer Durchbildung mehr und mehr Aufgaben übernehmen können, für die heute nur der Fachpsychologe die nötige Sachkenntnis und methodische Strenge besitzt. Und wenn dann einmal „schulpsychologische Ämter" entstehen werden, so werden sie nicht durch theoretische Wissenschaftler zu besetzen sein, sondern durch solche Pädagogen, die ein besonders hohes Maß psychologischer und jugendkundlicher Schulung mit gründlichen Unterrichts- und Erziehungserfahrungen verbinden[1]). Der gelehrte Fach-

[1]) Mit Nachdruck hat Otto Lipmann (*199, 200*) den Gedanken vertreten, daß für die Schülerauslese zwischen Psychologie und Praxis ein Vertrauensverhältnis anzubahnen sei, bei dem der gelehrte Fachpsychologe sich allmählich mehr und mehr entbehrlich machen müsse. — Zühlsdorff, der die psy-

psychologe wird noch immer genug damit zu tun haben, in seinen Forschungsinstituten die wissenschaftliche Vorarbeit in der Ausbildung der Methoden und die jugendkundliche Verarbeitung der Ergebnisse durchzuführen.

1. Die Vorauslese durch den Lehrer und die Abfassung von Schülerpsychogrammen.

Die erste Instanz für die vorzunehmende Auslese muß zweifellos der Lehrer sein, aus dessen Klasse der Schüler stammt. Dieser Satz wird zuweilen dahin zugespitzt, das der abgebende Lehrer die einzige Instanz zu bilden habe. Er kennt die Leistungen der Schüler aus monatelanger, zuweilen jahrelanger Schulbeobachtung und kann daher leicht zu der Meinung kommen, daß ihm auf Grund seiner praktischen Erfahrungen das alleinige Entscheidungsrecht für die Überführung zu einer höheren Bildung zustehe.

Hiergegen sind gewichtige Einwände zu erheben. So wertvoll seine Erfahrungen sein mögen, sie reichen allein nicht zur Entscheidung aus; er kann nur ein Vorschlagsrecht haben, also eine Vorauslese treffen. Aber auch für diese kann man sich nicht mit der bloßen Nennung begnügen; er muß vor sich selbst und den anderen beteiligten Instanzen seinen Vorschlag durch eine genaue psychologische Charakteristik der vorgeschlagenen Schüler begründen und dadurch die Entscheidung erleichtern.

Zunächst ist die Einstellung des durchschnittlichen Lehrers ganz naturgemäß eine „pädagogische" im engeren Sinne, d. h. auf den Erfolg des Unterrichts gerichtete. Er ist gewöhnt, den Schüler unter dem Gesichtspunkt der guten oder schlechten Leistungen zu betrachten; unter „Leistung" ist hier aber nicht das Werk schöpferischer, aus dem Innern quellender Selbsttätigkeit, sondern das äußerlich hervortretende Ergebnis der Unterrichtsbemühungen zu verstehen. Umfang und Korrektheit von Kenntnisleistungen und Fertigkeiten bestimmen gewöhnlich in erster Linie das Urteil der Lehrer über ihre Schüler. Bei der Schülerauslese aber, wie wir sie in Zukunft anstreben, lautet die Frage nicht mehr: ob ein bestimmtes Kind eben jetzt diejenigen Kenntnisse und Fertigkeiten besitze, die zum Mitkommen in einem anderen Schulzweig erfor-

---

chologischen Gesichtspunkte der Auslese durchaus würdigt, will doch den Psychologen möglichst ausschalten, vielmehr den Lehrer befähigen, die Auslese allein vorzunehmen.

derlich sind, sondern ob ihm die Fähigkeiten innewohnen, die es für eine höhere Ausbildung und entsprechende Berufswahl prädestinieren. Über diese Fähigkeiten aber fehlt dem durchschnittlichen Lehrer das genügend sichere Urteil, insbesondere wenn er sich nur auf seine Alltagsbeobachtung verlassen kann und keine Anleitung zu deren psychologischer Deutung erhält. Kenntnisse und Fertigkeiten des Schülers können aus den verschiedensten psychologischen Quellen entspringen: aus allgemeiner Intelligenz, irgend einer Sonderbegabung, gutem mechanischen Gedächtnis, Fleiß; auch häusliche Hilfe oder besondere Vorbildung sind an ihrem Zustandekommen in einer für den Lehrer meist unkontrollierbaren Weise beteiligt. Außer den Kenntnissen und Fertigkeiten aber gibt es noch zahlreiche andere Kriterien, aus denen man wichtige Aufschlüsse über Fähigkeiten der Kinder gewinnen könnte, wenn man sie nur beachtete (z. B. die Art des Spielens, der Verkehr mit den Kameraden, häusliche Beschäftigung usw.).

Gewiß gibt es Lehrer, die, mit angeborenem psychologischen Tiefblick begabt oder durch gründliche Beschäftigung mit der Psychologie geschult, imstande sind, sich so weit von den üblichen Einschätzungen der äußeren Leistungserfolge frei zu machen, daß sie zu einem wirklichen Bild der inneren psychischen Anlagen ihrer Schüler kommen; aber sie bilden doch nur Ausnahmen, auf die man bei der Organisation der Begabungsauslese nicht rechnen darf[1]). Es gibt andere Lehrer — und ihre Zahl wird nicht ganz klein sein —, die infolge ihrer rein didaktischen Einstellung die „Musterschüler" ohne weiteres für die Fähigsten halten; solche Musterschüler aber sind oft weit mehr durch Korrektheit, Bravheit und treffsicheres Gedächtnis als durch wirklich hohe Begabung ausgezeichnet; andererseits können Schüler, die infolge besonderer Veranlagung sich nicht ohne weiteres den üblichen Schulanforderungen einfügen und dem Lehrer etwas „unbequem" sind, leicht bezüglich ihres intellektuellen Wertes unterschätzt werden.

Noch ein anderes Bedenken spricht dagegen, dem abgebenden Lehrer die alleinige Entscheidung über die Auslese zu überlassen. Die Kinder, die in einer bestimmten Begabungsklasse vereinigt werden sollen, stammen aus sehr verschiedenen Abgangsschulen; bei ihrer

[1]) Wie verschieden die Fähigkeit der Lehrer ist, sich über die Fähigkeiten ihrer Schüler (unter Absehung von den äußeren Leistungserfolgen) Rechenschaft zu geben, ist bereits an anderer Stelle dieses Buches (S. 201 f.) ausgeführt worden.

Benennung werden die abgebenden Lehrer sehr verschieden strenge Maßstäbe anlegen, und die Folge ist eine große Ungleichmäßigkeit des Schülermaterials. Der eine, rigorosere Lehrer hält aus seiner Klasse nur ein einziges Kind für so begabt, daß es eine höhere Schulbildung verdient. Der Lehrer einer anderen Schule benennt fünf, nicht weil er so viel mehr begabte Schüler hat als jener, sondern weil er weitherziger ist, nicht so hohe Anforderungen stellt, sich von den Bitten der Eltern bestimmen läßt. Es wäre eine offenkundige Ungerechtigkeit gegen die Schuljugend, wollte man solche Entscheidungen für endgültig halten; es müßte dann so mancher Schüler der ersten Schule außerhalb der Auslese bleiben, der die Aufnahme vielleicht weit mehr verdient hätte, als mancher von dem weitherzigen Lehrer vorgeschlagene Schüler. Eine kontrollierende und ausgleichende Überprüfung erscheint unbedingt erforderlich.

Aber auch dann, wenn die abgebenden Lehrer lediglich die Vorauslese zu treffen haben, bleibt noch immer eine Gefahr bestehen. Die Vorauslese ist ja endgültig in negativer Hinsicht: d. h. diejenigen Schüler, die hier nicht genannt werden, kommen erst gar nicht in die engere Wahl. Ob hier nicht zuweilen wertvolle Begabungen von der höheren Ausbildung ausgeschlossen bleiben, weil der Lehrer sie übersehen oder nicht richtig eingeschätzt hat? Gegen diese Gefahr gibt es keinen unbedingten Schutz; sie kann aber beträchtlich vermindert werden durch folgende Maßregeln. Dem Lehrer wird nahegelegt, bei der Vorauslese so weitherzig wie möglich zu sein, selbst auf die Gefahr hin, daß eine größere Zahl der von ihm vorgeschlagenen Kinder nachher keine Aufnahme finden kann. Es ist dann immerhin die Wahrscheinlichkeit geringer, daß geeignete Kinder erst gar nicht zum Wettbewerb zugelassen werden. Der Lehrer mag dann unter den von ihm vorgeschlagenen Kindern diejenigen kennzeichnen, die er in erster Linie für die höhere Schulbahn empfiehlt. Sodann muß den Eltern bzw. (bei älteren Schülerjahrgängen) den Bewerbern selbst die Möglichkeit einer Berufung gegeben sein, wenn der abgebende Lehrer gegen ihren Willen den Schüler nicht vorschlägt. Von dieser Berufungsinstanz soll später die Rede sein.

Endlich muß die Vorauslese aufs gründlichste psychologisch gestützt und begründet sein; und hier stehen wir an der Stelle, wo die Psychologie ihre erste Mitarbeit zu gewähren hat.

Für jeden Schüler, der zur engeren Auslese kommt, muß ein Psychogramm, ein Seelenbild vorliegen, das, weit über die wenig besagenden Noten eines Schulzeugnisses hinaus, die wesentliche Richtung seines Geisteslebens und Grad und Art seiner Fähigkeiten erkennen läßt. Daß nur der Lehrer über die hierzu erforderlichen Erfahrungen verfügt, wenigstens potentiell, wurde schon angedeutet; aber er kann sie von sich aus nicht ohne weiteres nutzbar machen, sei es, daß seine praktisch-pädagogische Einstellung ihn vorwiegend auf die äußeren Leistungserfolge achten läßt (siehe oben S. 253), sei es, daß ihm die rechten Gesichtspunkte und Fragestellungen für die psychologische Beobachtung fehlen. Eben diese Gesichtspunkte und Fragestellungen vermag ihm nun die Psychologie, und nur diese zu geben. Sollen sich jene potentiellen, unbewußt und unbeachtet bleibenden Eindrücke in aktuelle Erfahrungen verwandeln und zu einer für die Vorberatung und Lebensbahnentscheidung des Kindes verwertbaren Darstellung verdichten, dann müssen Beobachtungsanleitungen ausgearbeitet und die Lehrer zu ihrer rechten Verwertung angeregt werden.

Die Forderung nach der Anlage psychologischer Personalbogen ist ja alt; sie ist auch bereits in einer Reihe von Sonderanstalten (z. B. heilpädagogischen Anstalten, Hilfsschulen) durchgeführt. Ihre Einführung in den Normalschulen scheiterte bisher an scheinbar unüberwindlichen Hindernissen. In der Tat würde die ohnehin schon stark belastete Zeit eines Lehrers von 40 bis 50 Schülern durch die Forderung, für jeden ein Psychogramm herzustellen, über die möglichen Grenzen hinaus beansprucht werden. Auch ist die Ausfüllung solcher Bogen für die Durchschnittsmasse der Kinder, die sich nicht durch Besonderheiten bemerkbar machen, außerordentlich schwer; die Bogen würden zum mindesten sehr wenig differenziert sein. Dagegen ist die Teilforderung der Ausfüllung solcher Bogen für die kleine Zahl der besonders befähigten Kinder durchführbar und hier auch leichter, weil diese Kinder durch ihre Gaben auffallen und sich gleichsam von selbst individualisieren. Für diese aber ist die Forderung auch zur Zeit am dringlichsten, weil eben diese Kinder die Anwärter für besondere Umschulungen und Einschulungen sind.

Die Abfassung von Schülerpsychogrammen stößt nun auf eine ganze Reihe von Schwierigkeiten, die erst bei der praktischen Arbeit — wir verfügen in Hamburg bereits über 3000 solcher ausgefüllten Beobachtungsbogen — in ihrer ganzen Schärfe hervortreten.

Der Beobachtungsbogen muß in der Anlage den psychologischen Anforderungen entsprechen, die auf Feststellung der wichtigsten psychischen Dispositionen des Kindes gehen; er muß andererseits dem Verständnis und dem Interesse des ausfüllenden Lehrers in Ausdrucksweise und Umfang angemessen sein. Daraus ergibt sich die Notwendigkeit gewisser Kompromisse. Hält man sich streng an die Einteilung und Ausdrucksweise eines wissenschaftlich psychologischen Systems, so wird der praktische Zweck des Bogens illusorisch; ich habe so manche Entwürfe von Bogen gesehen, die nur zu deutlich ihren Ursprung vom Schreibtisch des schulfremden Psychologen verrieten. Ist das psychographische Schema zu lang, so wird der Lehrer abgeschreckt; gehen die Fragen zu sehr ins einzelne, so empfindet der Lehrer sie als Zwangsroute und als Suggestion. Macht man aber den Bogen andererseits zu kurz, so sind die dargebotenen allgemeinen Begriffe so unbestimmt und vieldeutig, daß der Ausfüller entweder überhaupt nichts mit ihnen anzufangen weiß oder sie in verschiedenster Weise mißdeuten kann. Hier wird man vor allem das oft zu hörende Vorurteil des Praktikers bekämpfen müssen, daß kürzere psychographische Schemata leichter zu bewältigen seien als lange. Das mag für Fragelisten über körperliche und sonstige äußerlich feststellbare Merkmale gelten, für psychologische Beobachtungsanleitungen trifft es nicht zu. Wird z. B. nur kurz nach der „Intelligenz" gefragt, so ist der Ausfüller oft genug ratlos; wird dagegen jenes allgemeine Stichwort in eine Reihe von Spezialfragen aufgelöst (z. B. Neigung des Kindes zu selbständigem Fragen, Vorauseilen mit seinem Denken bei Behandlung des Unterrichtsstoffes, Fähigkeit das Wesentliche zu erfassen, Neigung zur Kritik usw.), so kann sich der Lehrer viel mehr in dem ihm vertrauten Gebiet konkreter Erfahrungen bewegen, und das Schema gewinnt Leben und Farbe für ihn. Je kürzer und allgemeiner die Liste der Gesichtspunkte ist — um so ausführlicher müssen die Erläuterungen sein, die zum Verständnis und zur praktischen Verwertung der Liste unentbehrlich sind.

Vielfach wendet sich der Widerstand der Lehrerschaft gegen das Prinzip eines psychographischen Beobachtungsschemas überhaupt. Es wird anerkannt, daß bei den Schülern, deren Schulbahnberatung in Frage steht, eine wirkliche Berücksichtigung der Individualität stattfinden müsse, und daß zu diesem Zweck auch die Darstellung eines Seelenbildes zu erfolgen habe. Aber dies

Seelenbild soll nicht als Mosaikbild an der Hand einer Liste, sondern als freie abgerundete „Charakteristik" vom Lehrer gegeben werden. Die Liste zerreiße die Persönlichkeit in eine Reihe von Einzelfunktionen; die freie Charakteristik solle die Wesenheit jedes Kindes aus dem einheitlichen Quellpunkt seiner Persönlichkeit heraus schildern, ohne durch Vorschriften an Gesichtspunkte gebunden zu sein, die für dieses Kind vielleicht unwesentlich sind.

Wer wie der Schreiber dieser Zeilen im Personalismus die Grundlage seiner wissenschaftlichen Überzeugung sieht, der ist wohl gegen den Verdacht gefeit, die Persönlichkeit des Kindes in der Personalbeschreibung atomisieren und „versächlichen" zu wollen. Für ihn ist die Persönlichkeit „unitas", ein unteilbares, auf die Einheit innerer Zweckbestimmtheit eingestelltes Wesen. Aber er bekämpft zugleich den „naiven" Personalismus, der diese Einheit zur Einfachheit verengt, der einen einzigen Zug gleichsam zur Quintessenz zu erheben sucht. Die Person ist „unitas multiplex", nicht „unitas simplex". Die Gefahr einer solchen Versimpelung und Vergröberung des Schülerbildes ist nicht minder groß als die der Atomisierung; der Versuch, die Wesenheit eines Kindes auf eine einfache Formel zu bringen, ist nur allzu oft das Erzeugnis apriorischer Konstruktion oder einer unkontrollierbaren Intuition des Lehrers — bei der womöglich dessen eigener psychologischer Typ Gevatter gestanden hat. Die Zahl der Lehrer, die zur Darstellung künstlerisch abgerundeter Vollbilder ihrer Schüler befähigt sind, wird stets sehr klein bleiben; die überwiegende Mehrzahl wird sich auf dürftige Umrißlinien beschränken, die die Fülle der in der Persönlichkeit vorhandenen Farben und Schattierungen nicht ahnen lassen; das Schülerbild wird viel schematischer ausfallen, als wenn ein psychographisches Schema benutzt worden wäre.

Der Beobachtungsbogen darf aber nicht allein vom Standpunkt des einzelnen Kindes aus beurteilt werden, sondern auch vom Standpunkt des notwendigen Vergleiches vieler Schüler. Bei der Auslese müssen ja Kinder verschiedenster Schulherkunft gegeneinander abgewogen werden; liegen über diese lediglich freie Beschreibungen aus der Feder verschiedener Lehrer vor, so fehlt jede Möglichkeit einer wirklichen Vergleichung. Der eine Lehrer nennt diesen, der andere jenen Zug; der erste läßt sich ausführlicher über das Gedächtnis aus und sagt nichts über Interessen und über Denkfähigkeit; ein zweiter wiederum spricht vornehmlich von der Phantasie,

ein dritter von der Intelligenz unter Vernachlässigung der übrigen Funktionen. Ganz anders, wenn eine Reihe von Gesichtspunkten vorliegt, durch die jeder Ausfüller veranlaßt wird, auf bestimmte Hauptzüge zu achten. Sind dann einzelne Punkte nicht ausgefüllt, so wissen wir, daß dies nicht auf Zufall oder Vergeßlichkeit beruht, sondern darauf, daß der Lehrer hier nichts Bemerkenswertes zu beobachten vermochte.

Selbstverständlich darf das Schema nicht zu einer Zwangsjacke werden. Es soll nicht imperativische Bedeutung haben, als ob jede Frage nun unbedingt beantwortet werden müsse; im Gegenteil, die begleitende Anweisung muß unzweideutig den Satz enthalten, daß nur diejenigen Rubriken auszufüllen seien, über welche sichere und bedeutsame Beobachtungen vorliegen. Ebensowenig soll der Lehrer glauben, daß er nur bringen dürfe, was der Bogen erfragt. Wenn er Beobachtungen hinzufügt, die durch den Bogen nicht nahegelegt waren, aber für das Kind kennzeichnend sind, wenn er Proben kindlicher Betätigung (z. B. Zeichnungen, Aufsätze, Briefe usw.) beilegt, wenn er endlich auf Grund der Bausteine, die der Bogen liefert, eine zusammenfassende abgerundete Charakteristik anfügt, die nun die leitenden Züge der kindlichen Persönlichkeit herausarbeitet, dann werden jene Forderungen der Bogengegner erfüllt, ohne daß die Vorzüge des Bogens geopfert werden müßten.

Für die Anlage des Beobachtungsbogens ist bisher noch nicht die endgültig befriedigende Form gefunden. Zwischen einem eigentlichen Fragebogen, der nach möglichst konkreten Beobachtungen fragt, und einem bloßen Berichtsplan, der die Hauptgesichtspunkte in Stichwörtern gibt und dem Ausfüller ganz die nähere Detaillierung überläßt, schwanken die Vorschläge noch hin und her. Die Form der Befragung birgt, wie schon erwähnt, die Gefahr starker Suggestion und Einengung der Beobachtung in sich; die Stichwortform kann allzuleicht durch ihre Vagheit zu dürftigen, unklaren und wenig vergleichbaren Darstellungen führen. Als Verstärkung der Suggestion hat es augenscheinlich gewirkt, wenn man den einzelnen Fragen eine Reihe möglicher Antworten beifügte, z. B.: „Stellt das Kind selbständig sinnvolle Fragen? (Ja; häufig; X. versucht stets, der Sache auf den Grund zu kommen; neigt zum Weiterdenken; selten)"; es zeigte sich nämlich, daß viele Ausfüller glaubten, sich wörtlich an diese Antwortliste halten zu müssen und auf eigene Formulierung ihrer Beobachtungen verzichten zu können. Bewährt hat sich dagegen die Beigabe von „Beobachtungsgelegenheiten" z. B.: „Ist das Kind phantasiebegabt oder nüchtern? (zu beobachten: beim Aufsatz; beim Spiel; im Zeichenunterricht; beim Wiedererzählen; bei Spielereien in der Stunde; an der häuslichen Beschäftigung)"; denn hierdurch werden die Ausfüller auf manche Möglichkeiten der Beobachtung hingewiesen, die ihnen sonst entgangen wären.

Insbesondere ist es wichtig, daß sie veranlaßt werden, über den Rahmen des bloßen Unterrichtsbetriebes hinauszusehen und auch Beobachtungen aus dem Verhalten des Kindes beim Spiel, in kameradschaftlichem Verkehr, im Hause heranzuziehen; denn hier liegen oft charakteristischere Bekundungen der kindlichen Eigenart als in dem Verhalten beim Lernen und Arbeiten. Um diese Abrundung des Bildes zu ermöglichen, ist Verbindung und Gedankenaustausch mit dem Elternhause dringend zu empfehlen.

Notwenig erscheint ferner der Hinweis, daß der Lehrer Beobachtungen, nicht aber Tests anstellen solle. Denn das Bild, das er gibt, soll ja gerade das spontane, natürliche Verhalten des Kindes schildern, nicht seine Reaktion auf gewisse künstliche Anforderungen; es soll die pädagogische und psychologische Prüfung aufs wertvollste ergänzen, nicht aber sie in unkontrollierbarer Weise einbeziehen. Die dilettantische Verwertung von Tests durch Unberufene und Ungeübte kann viel mehr Schaden als Nutzen stiften.

Zur Technik der Ausfüllung ist vor allem zu bemerken, daß der Lehrer rechtzeitig den Bogen in Händen haben muß. Je längere Zeit ihm zur Verfügung steht, um zwanglos Beobachtungen anzustellen und zu notieren, um so größer ist die Gewähr, daß die Ausfüllung reichhaltig und zuverlässig wird. Wird der Bogen wenige Wochen vor dem Ablieferungstermin dem Lehrer eingehändigt, so ist zu befürchten, daß die Ausfüllung stark auf Vermutungen, unbestimmte Erinnerungen und mehr oder minder zufällige Einzelbeobachtungen gegründet sein wird. Am empfehlenswertesten ist es, wenn dem Lehrer dauernd eine Beobachtungsanleitung zur Verfügung steht und wenn er auf Grund dieser für jeden Schüler, der sich durch besondere Fähigkeiten bemerkbar zu machen beginnt, eine Liste oder ein Beobachtungsheft anlegt, woraus dann später das endgültige Psychogramm abgeleitet werden kann[1]).

Ein unbedingtes Erfordernis sind geeignete Maßnahmen, durch die der Lehrer in Anlage und Benutzungsart des Beobachtungsbogens eingeführt wird. Hierzu dienen zunächst gedruckte Anweisungen und Erläuterungen, die jedem Ausfüller in die Hand gegeben werden; sie enthalten einerseits allgemeine Richtlinien, sodann nähere Erklärungen der einzelnen Abteilungen des Bogens. Und zwar müssen, wie schon gesagt, diese Erläuterungen um so umfangreicher sein, je knapper im Bogen selbst die Formulierungen der Gesichtspunkte sind. Ferner sind mündliche Aufklärungen sehr empfehlenswert, sowohl durch Vorträge für alle beteiligten Lehrer über Bedeutung und Inhalt des Bogens, wie durch psychologisch-pädagogische Sprechstunden, in denen die Lehrer sich über Zweifelsfragen Orientierung holen können. Weiter ausschauend sind die Forderungen, die auf eine ganz andere Schulung der

[1]) Es hat dies auch den technischen Vorzug, daß nicht für jeden vorzuschlagenden Schüler, sondern nur für jeden vorschlagenden Lehrer ein Druckstück des Beobachtungsbogens geliefert zu werden braucht.

Lehrerschaft in Kindes- und Jugendpsychologie gehen; die Erfahrungen, die bisher mit Beobachtungsbogen gemacht worden sind, sprechen gerade in dieser Hinsicht eine beredte Sprache. Die Ratlosigkeit, mit der viele Lehrer und Lehrerinnen vor der Aufgabe der Psychographie ihrer Schüler stehen, beruht meiner Überzeugung nach nur zum kleineren Teil auf Unvollkommenheit des Verfahrens, zum größeren Teil auf der Fremdheit gegenüber wirklich psychologischer Erfassung und Durchdringung der kindlichen Psyche. Und dies kann natürlich nicht durch einzelne Vorträge und gedruckte Anweisungen, sondern nur dadurch gründlich geändert werden, daß die Kindes- und Jugendpsychologie in der Vorbildung und Fortbildung des Lehrers einen Platz erhält, der ihrer erhöhten Bedeutung im künftigen Kulturleben entspricht.

Folgende Veröffentlichungen von und über Beobachtungsbogen zur Schülerauslese liegen bisher vor: Die Arbeitsgemeinschaft für exakte Pädagogik in Berlin ließ einen sehr ausführlichen Bogen durch Rebhuhn veröffentlichen (*221*), der auch im Einzeldruck zu kaufen ist. Einen Breslauer Bogen entwarf Alfred Mann (*203*). Der für 10jährige bestimmte Hamburger Bogen, bearbeitet von Martha Muchow, liegt bereits in 2 Fassungen vor, deren zweite auch als Sonderdruck erschien (*211*). Über die Erfahrungen mit diesem Bogen berichten die Arbeiten von Martha Muchow (*213, 218*) und von Muchow und Höper (*214*). In Anlehnung an den Hamburger Bogen hat das österreichische Unterrichtsministerium das Schema einer „Schülerbeschreibung" herausgegeben, das der Aufnahme in die Unterklasse der Mittelschulen zugrunde gelegt werden soll (*227, 245*); das Schema ist auch abgedruckt in ZAngPs (*237*). Ebenfalls im Anschluß an den Hamburger Bogen ist der Stettiner ausgearbeitet, den Zühlsdorff (*244*) veröffenttlicht. Beobachtungslisten aus Leipzig und Frankfurt a. M. erscheinen im 21. Bande der ZPdPs.

Der Entwurf zu einem Individualitätsbogen zur Benutzung bei der Begabtenauslese im vierten Schuljahr enthält auch das Buch von Schneider über „Das Studium der Individualität" (*222*a). — Weitere Literatur, namentlich auch über Beobachtungsbogen für Berufsberatungszwecke, ist zusammengestellt in einem kleinen Sammelbericht von Lipmann (ZAngPs **17**, S. 195 f.).

2. Die Prüfung.

Überall wo die Vorauslese der abgebenden Schulen mehr Kinder benennt als aufgenommen werden können, muß eine engere Auslese durch ein Prüfungsverfahren vorgenommen werden. Das gleiche gilt für diejenigen Fälle, bei denen die Meldungen zu einer bestimmten Schulbahn frei (ohne Vorauslese) erfolgen können, z. B. bei dem Zugang zum Lehrerseminar, zu einer Werkschule und anderen Fachschulen. Endlich sind Prüfungen dort unentbehrlich,

wo jemand auf einem anderen als dem üblichen Wege zu einer Vorbildungsmöglichkeit (z. B. zur Hochschule) kommen möchte.

Es ist hier nicht der Ort, die schwere Frage einer Reform des Prüfungswesens überhaupt anzuschneiden. Der Wunsch, die Prüfungen ganz abzuschaffen, wird leider stets an der eisernen Härte der Verhältnisse scheitern; dagegen ist eine weitgehende Verbesserung und Umgestaltung des Prüfungswesens gerade durch das Vorbild der Begabtenauslese zu erhoffen. Diese liefert nämlich zwei neue Gesichtspunkte: sie stellt erstens der üblichen Fertigkeits- und Kenntnisprüfung die „Fähigkeitsprüfung" zur Seite; sie fordert zweitens, daß eine Prüfung allein nicht entscheidend sein, sondern nur als ein Kriterium neben anderen bei der Entscheidung mitwirken darf.

Der Gegensatz zwischen einer rein pädagogischen und einer psychologischen Einstellung, der uns bei den abgebenden Lehrern schon beschäftigte, begegnet uns auch bei den Lehrern der aufnehmenden Schulen. Die gewöhnlichen Aufnahmeexamina haben einen durchaus schultechnischen Zweck: sie sollen in erster Linie feststellen, ob bei dem Bewerber die stofflichen Voraussetzungen vorhanden sind, auf denen dann das Pensum der neuen Klasse weiter zu bauen hat; es werden deshalb gewisse Fertigkeiten und Kenntnisse geprüft (z. B. im Rechnen und Diktat beim Übergang von der Grundschule zur untersten Klasse der höheren Schule). Die Frage, ob die Prüflinge überhaupt die Eignung zum Besuch der höheren Schule haben — die weitergehende, ob die besondere Begabung des Kindes gerade den Besuch dieser Schule nahelege — und die weitestgehende: ob die Fähigkeiten des Kindes dem Staat nicht geradezu die Verpflichtung auferlegen, ihm den Zugang zu einer höheren Bildung bestimmter Art zu öffnen — sie liegen den üblichen Aufnahmeprüfungen mehr oder weniger fern.

Aber gerade diese Fragen sind es, die in Zukunft durchaus im Vordergrunde zu stehen haben; und darum müssen wir grundsätzlich zwischen Aufnahmeprüfungen und Ausleseprüfungen — oder anders ausgedrückt: zwischen Fertigkeits- und Fähigkeitsprüfungen — scheiden. Dabei soll nicht etwa einer völligen Beseitigung der Fertigkeitsprüfung das Wort geredet werden. Zu der Möglichkeit, an dem Unterricht einer Klasse mit Erfolg teilzunehmen, gehört nun einmal ein bestimmtes Maß vorgängig erworbener Kenntnisse und Fertigkeiten; und auch psychologisch ist der Nachweis nicht ganz unwichtig, ob der Prüfling diejenige

Lern- und Übungsfähigkeit besitze, wie sie sich in dem Besitz bestimmter Kenntnisse und Fertigkeiten bekundet. Niemals aber sollte eine solche Prüfung allein ausschlaggebend sein dürfen. Die fehlenden stofflichen Voraussetzungen zum Mitkommen in der Klasse können, wenn die Begabung des Kindes hoch ist, durch Nachhilfe und Sonderkurse nachgeholt werden; psychologisch aber gibt die einseitige Feststellung der Gedächtnis- und Übungsfähigkeit ein leicht irreführendes Bild von dem Stand der geistigen Gesamtfähigkeit des Kindes.

Deshalb sind zum mindesten daneben eigentliche Fähigkeitsprüfungen nötig, die nicht auf eingeübte Fertigkeiten und erworbene Kenntnisse, sondern gerade auf Nichterlerntes und Nichterlernbares, auf die Höhe der geistigen Anlage und auf deren vermutliche künftige Entwicklung gerichtet sind. Und zu deren Ausgestaltung und Anwendung ist wiederum die Psychologie nötig, die ihre Methoden der Begabungstests nun den neuen Forderungen gemäß auszugestalten hat.

Je nach den besonderen pädagogischen Gesichtspunkten der Auslese wird die Einbeziehung der Testprüfung sehr verschiedene Formen annehmen können; deshalb werden sich nur wenige allgemeingültige Grundsätze aufstellen lassen.

Für die Auswahl der anzuwendenden Tests sind folgende Bedingungen zu berücksichtigen:

a) Die Schulgattung, für welche ausgelesen werden soll. Jede Schule, die sich aus der allgemeinen Grundschule heraushebt, verlangt von ihren Zöglingen zunächst einen gewissen, den Durchschnitt überschreitenden Grad allgemeiner geistiger Leistungshöhe; deshalb werden manche Tests, welche Auffassung und Verständnis, Kombinationsfähigkeit, Gabe der Kritik, das Sich-Abfinden mit ungewohnten Aufgaben usw. prüfen, für alle Auslesen in Betracht kommen. Außerdem aber sind die einzelnen Schulen nach den besonderen Aufgaben — und infolgedessen auch nach den für sie vor allem wichtigen Befähigungsrichtungen — unterschieden; diesen Unterschieden muß die jeweilige Testauswahl ebenfalls gerecht werden. Für eine humanistisch-sprachliche Schule muß die sprachliche Seite der Intelligenz, für die Werkschule eines technischen Betriebes muß die konstruktive Fähigkeit und praktische Intelligenz, für ein Lehrerinnenseminar müssen die pädagogischen Seiten der Intelligenz zur Feststellung gelangen. Bezüglich solcher

Differenzierungen ist das meiste erst von der Zukunft zu erwarten; bisher hat man sich von psychologischer Seite hauptsächlich auf die Ausarbeitung solcher Methoden beschränkt, welche die neutrale Allgemeinintelligenz zu untersuchen geeignet sind.

b) Alter und Fähigkeit der Prüflinge. Die Tests müssen die charakteristische Richtung und Einstellung der geistigen Arbeit treffen, die dem geprüften Alter eigen ist; sie müssen zugleich so gestaltet sein, daß die Ergebnisse zu einer deutlichen Differenzierung der Gleichalterigen führen und somit die Begabtesten erkennbar hervortreten lassen. Der Fehler liegt hier nahe, daß man „bewährte" Tests auf die neuartigen Bedingungen der Begabungsprüfung unbesehen überträgt und dadurch zu unklaren oder gar irrigen Ergebnissen kommt. Es gibt in diesem Sinne keine „bewährten Tests überhaupt", sondern nur solche, die für bestimmte Zwecke und Altersstufen geeignet sind. Die an sich so wertvollen Binet-Simon-Tests z. B. haben ihr Anwendungsgebiet ganz wesentlich bei normalen und unternormalen Kindern; bei begabten versagen sie leicht, weil die Ansprüche, die sie stellen, zu einfach sind, zu wenig höhere Denkarbeit verlangen; und manche Tests, die sich für 10jährige vorzüglich eignen, sind, selbst in erschwerter Form, für 14jährige wenig charakteristisch, weil sich die entscheidenden Leistungsformen dieses Alters in ganz anderen geistigen Richtungen bewegen[1]).

Die wissenschaftliche Vorbereitung der Begabungsprüfungen hat sich daher in den letzten Jahren mit Recht darauf konzentriert, „höhere" Intelligenztests auszuarbeiten, d. h. eben solche, die dem gesteigerten Grade der Denkfähigkeit, der größeren Selbständigkeit und dem vertieften Verständnis der Begabten angemessen sind; so sind die Kritiktests, die Ordnungstests, die Schematests, die Bilderbogentests, die neuen Formen der Lückenergänzungen und der Dreiwortmethoden, die Beurteilungstests usw. wesentlich aus diesem Bedürfnis heraus ausgebildet worden. Ferner muß für alle anzuwendenden Tests aufs dringlichste eine vorgängige Eichung gefordert werden: an einem größeren Material normaler Kinder ist festzustellen, ob der Test selbst, die Form seiner Anwendung, der Wortlaut der den Kindern zu gebenden Anleitung brauchbar ist, und ob er gerade für dieses Alter die erforder-

[1]) Hier muß auf jene qualitativen Wandlungen der Intelligenz Rücksicht genommen werden, von denen wir früher (S. 33) sprachen.

lichen deutlichen Differenzierungen der Begabung liefert. Ohne solche vorgängige Eichung kann es bei der Prüfung selbst die peinlichsten Überraschungen geben[1]).

Für diese Eichung sind verschiedene Wege gangbar.

So kann man die im Kapitel „Intelligenz und Schulleistung" besprochene Korrelationsmethodik anwenden; man stellt an den Kindern entsprechender Klassen fest, in welchem Maße das aus dem Test zu gewinnende Fähigkeitsbild mit der von den Lehrern herrührenden Intelligenzschätzung oder psychographischen Beschreibung der Prüflinge sich deckt. Auf diesem Wege hat z. B. Susanne Engelmann (*187*a) den Nachweis geführt, daß gewisse Tests ein recht zutreffendes Bild der Befähigungen liefern und daher mit Fug der Begabungsprüfung zugrunde gelegt werden dürfen. Auch die anderen in dem genannten Kapitel erwähnten Ergebnisse über die Korrelation von Intelligenzschätzung und Testprüfung sind in gleichem Sinne zu verwerten.

Ein anderer Weg zur Eichung besteht in Vorversuchen an Kindern mehrerer Altersstufen, und zwar am besten bei der Altersstufe, die den Ausleseprüflingen entspricht, und den beiden nächsthöheren Jahrgängen. Zeigt der Test in diesen drei Altersstufen einen starken Altersfortschritt und innerhalb jeder Altersgruppe gute Korrelation zur Begabung, so ist er recht brauchbar; denn man kann im allgemeinen von den Begabtesten einer Altersstufe annehmen, daß sie bezüglich ihrer Intelligenzleistung mit den um etwa zwei Jahre älteren normalen Kindern auf eine Stufe zu stellen sind.

Daß solche Eichungs-Vorversuche den künftigen Prüfungsanwärtern völlig verborgen bleiben müssen, ist selbstverständlich. Alle Beteiligten müssen strengste Verschwiegenheit bewahren; um ein Herumsprechen unter den Kindern selbst zu verhindern, empfiehlt es sich, die Vorversuche an einem anderen Ort anzustellen. (So wurden die Tests für Hamburg in Wandsbek oder Bergedorf geeicht.)

c) Die Anwendung im Massen- oder Einzelversuch. Aufs

---

[1]) Es genügt auch nicht etwa, aus der Methodensammlung von Stern und Wiegmann (*8, III*) oder aus dem Buch von Moede und Piorkowski (*206*) oder aus den Leipziger Anweisungen (*197*) Tests auf Grund der dort gegebenen Altersbestimmungen auszuwählen, und ohne Vorproben zur Auslese zu verwenden. Jeder Ort, jede Schulart, jede Situation hat ihre besonderen Bedingungen, denen die Tests angepaßt werden müssen — was nur durch eigene Vorversuche möglich ist.

stärkste werden stets die äußeren Verhältnisse der Auslese auf die Auswahl der Tests einwirken. Handelt es sich, wie in Berlin oder Hamburg, um die gleichzeitige Prüfung von mehreren hundert oder gar tausend Schülern, dann ist ja von vornherein die individuelle Prüfung ausgeschlossen, und man muß zum schriftlichen Verfahren greifen. Aber auch bei Auslesen kleineren Umfangs wird zuweilen die Gleichzeitigkeit des schriftlichen Arbeitens aus diesem oder jenem Grunde notwendig werden. Derartige Massenprüfungen haben ihre Nachteile und Vorteile.

Folgende Nachteile liegen vor. Die psychologische Beurteilung ist lediglich angewiesen auf das in Form einer Niederschrift vorliegende Endergebnis der Arbeit am Test; dagegen sind die einzelnen Phasen dieser Arbeit selbst für die Beobachtung und nachfolgende Beurteilung verloren; und doch sind diese Arbeitsweisen für Art und Grad der Befähigung zuweilen kennzeichnender als die größere oder geringere Güte des Enderfolgs. Infolgedessen entstehen auch bei der Deutung der schriftlichen Ergebnisse oft Schwierigkeiten, über die im Abschnitt „Verarbeitung" zu sprechen sein wird. Ferner ist die Auswahl der Tests dadurch beträchtlich eingeschränkt, daß nur solche mit schriftlicher Ausführbarkeit in Frage kommen; insbesondere sind „Handlungstests" so gut wie ausgeschlossen. Endlich ist beim Massenversuch jedes individualisierende Eingehen auf die besonderen Verhaltungsweisen der einzelnen Prüflinge unmöglich.

Aber eben diese Gleichmäßigkeit in der Behandlung aller Prüflinge macht auch die Stärke der Massenprüfung aus. Denn sie ermöglicht eine weitgehende Vergleichung der Prüflinge und die Herstellung einer exakten Rangordnung unter ihnen. Und wenn auch diese Rangordnung niemals zur alleinigen Richtlinie der Auslese gemacht werden darf, muß sie doch als ein Faktor der Beurteilung neben den anderen gebührend eingeschätzt werden. Die Massenprüfung zeigt uns die reaktive Leistungsfähigkeit der Kinder gleichsam in Reinkultur; die äußeren Bedingungen ihrer Arbeit: die Aufgaben, die Anleitungen dazu, die zur Verfügung stehenden Zeiten sind für alle Kinder übereinstimmend, und die Verschiedenheiten der Leistungen sind daher so gut wie restlos auf die inneren im Prüfling selbst liegenden Reaktionsbedingungen zurückzuführen.

Das Ideal dieser Gleichmäßigkeit ist natürlich niemals voll zu erreichen: aber es muß bei allen Maßnahmen der Massenprüfung vorschweben. Eine Grenze liegt schon in der Zahl der Prüflinge:

übersteigt diese die Ziffer von etwa 30, dann ist es schon nicht mehr möglich, alle gleichzeitig im selben Raum durch einen Prüfer prüfen zu lassen. In solchen Fällen muß man gruppenweise prüfen, und nun ist peinlich dafür zu sorgen, daß alle Gruppen gleich behandelt werden.

Hierzu gibt es zwei Möglichkeiten: entweder werden die Prüfungen von demselben Prüfer zu verschiedenen Zeiten vorgenommen; oder die Gruppen werden von verschiedenen Prüfern möglichst gleichzeitig geprüft. Da die erste Methode die Gefahr eines Gedankenaustausches zwischen schon geprüften und noch zu prüfenden Kindern mit sich führt, so wird meist der zweite Ausweg gewählt werden müssen. Bei ihm bringt nun freilich wieder die Verschiedenheit der Prüfer einen Ungleichheitsfaktor in das Verfahren; er muß dadurch auf ein Mindestmaß gebracht werden, daß die Prüfer in gemeinsamen Vorbesprechungen auf alle Einzelheiten des Verfahrens eingeübt werden. Am besten wählt man hierzu solche Persönlichkeiten, die mit den Gesichtspunkten der Jugendpsychologie und den Bedingungen der Begabungsauslese einigermaßen vertraut sind. Sie müssen genau ausgearbeitete Anweisungen erhalten, in denen jedes Wort, jeder Handgriff, jede Zeitbestimmung angegeben ist; sie müssen angewiesen sein, jede individuelle Maßnahme und Äußerung zu unterlassen, die vielleicht durch Zwischenfragen der Kinder und ähnliches nahegelegt sein könnte; sie haben über den Verlauf der Prüfungen genau Protokoll zu führen. Sind diese Prüfer praktische Pädagogen, so ist ihnen klarzumachen, daß pädagogisch wünschenswerte Methoden — z. B. daß man die Selbstbetätigung des Kindes in freier Weise spielen lasse, oder daß man durch Erläuterungen, Hilfsfragen usw. die Denktätigkeit des Kindes unterstütze — bei dieser psychologischen Feststellung unzulässig seien, weil alles auf die möglichst vollständige Gleichheit der Prüfungsbestimmungen ankomme[1]).

Die Massenprüfung erhält hierdurch unleugbar einen etwas mechanischen Charakter; und eben deshalb wäre es auch verfehlt, auf sie allein Entscheidungen zu gründen; aber da ja nun nach unseren Forderungen die anderen mehr individualisierenden Verfahrungsweisen neben ihr einen Platz haben sollen, so ist es angebracht, diese Teilmethode nach ihrem Vorzug, dem einer möglichst weitgehenden Vergleichbarkeit der Einzelleistungen, voll auszunützen.

Nicht minder peinlich wie bei der Prüfung selbst muß auch bei der Bearbeitung und Bewertung der Ergebnisse der Grund-

[1]) Übrigens wird selbst bei genauer Innehaltung der allgemeinen Vorschriften durch die einzelnen Prüfer niemals eine ideale Gleichmäßigkeit des Verfahrens in allen Gruppen erreicht werden. Denn letzte Persönlichkeitseigenschaften der Prüfer, wie Tonfall, Strenge oder Freundlichkeit des Wesens, Sicherheit oder Unsicherheit des Auftretens usw. entziehen sich nun einmal der Gleichmacherei.

satz der Gleichmäßigkeit gewahrt werden. Da gewöhnlich zwischen der Prüfung und der zu fällenden Entscheidung nur eine kurze Frist liegt, so muß die Durcharbeitung des Materials oft im Geschwindschritt erfolgen. Deshalb wird — namentlich wenn es sich um größere Prüflingsziffern handelt — auch hier Arbeitsvereinigung und Arbeitsteilung unvermeidlich sein.

Zunächst müssen für jeden benutzten Test Bewertungsregeln aufgestellt werden, die von allen Beteiligten streng innezuhalten sind. In Zweifelsfällen ist in gemeinsamer Besprechung eine Vereinbarung zu treffen. Schließlich sind alle gegebenen Prädikate von der die Bewertungsarbeit leitenden Persönlichkeit noch einmal zu überprüfen. Die Möglichkeit, solche gleichmäßigen Bewertungsgesichtspunkte aufzustellen und anzuwenden, ist nun bei verschiedenen Tests sehr verschieden. Es gibt manche Testformen — z. B. Aufsätze —, die für wichtige Seiten der Individualität sehr kennzeichnend sein mögen, deren Beurteilung aber von subjektiven Auffassungen des Beurteilers ziemlich weitgehend abhängig bleibt. Solche Tests sollte man deshalb bei Massenprüfungen nur sehr zurückhaltend verwenden.

Wichtig bei der Bewertung ist es, daß man sich stets die psychologische Natur des Tests vor Augen hält. Es wäre ganz verfehlt, wollte man mit rein schulmeisterlicher Einstellung an die vorliegenden Arbeiten herangehen und etwa nur nach der Anzahl der anzustreichenden Fehler das Urteil fällen. Vielmehr kommt bei den gemachten Fehlern alles darauf an, auf welche psychischen Momente sie zurückzuführen sind; und außer den Fehlern muß auch die positive Seite der Leistung, die verschiedenartige Weise, in der den Anforderungen entsprochen oder über diese gar hinausgegangen worden ist, berücksichtigt werden. Eine solche psychologische Rückdeutung ist, wie schon oben angedeutet, dadurch erschwert, daß der Beurteiler nicht den Arbeitsvorgang des Kindes selbst beobachtete, sondern nur die Endleistung vor sich sieht. Um diese Schwierigkeit wenigstens zu verringern, ist zweierlei zu fordern: es sind möglichst nur solche Tests anzuwenden, bei denen keine allzu große Vieldeutigkeit der Endleistung vorliegt; und es muß zweitens jeder Beurteiler Gelegenheit haben, die Anwendung des zu bewertenden Testes auch einmal in besonderen Einzelversuchen kennen zu lernen, hierbei die Kinder während der Arbeit zu beobachten und nach der Arbeit über Einzelheiten ihres Tuns zu befragen. Eine derartige Kontrolle erhellt oft mit einem Schlage Unverständlichkeiten und Dunkelheiten, die bei den Niederschriften begegnen.

Man ersieht aus diesen Andeutungen, daß die Bewertung der

Tests nicht minder psychologische Schulung voraussetzt wie die Anstellung der Prüfung selbst; wir müssen daher die hier und da laut gewordene Absicht zurückweisen, daß die Psychologie zwar bei der Ausarbeitung und Vorbereitung der Tests mitarbeiten, die Anwendung und Prädikatgebung aber dem praktischen Pädagogen allein überlassen solle. Die Ausarbeitung und Überwachung des Bewertungssystems muß in den Händen von Persönlichkeiten liegen, die durchgebildete Psychologen und insbesondere Kenner der begabungspsychologischen Methoden sind, die zugleich aber auch mit der pädagogischen Praxis vertraut sind.

Das Ergebnis einer solchen Bearbeitung besteht nun darin, daß jeder Schüler für jeden Test eine Wertziffer erhält. Diese Wertziffern sind untereinander zunächst sehr ungleichartig: bei einem Test gibt sie die Anzahl der Fehler, bei einem anderen die Schätzung der Leistungsgüte an; bei einem dritten (z. B. einem Aufsatz) besteht sie schon aus der Vereinigung verschiedener Teilwerte (z. B. für Stil, Phantasie, Reflexion, Beobachtungsfähigkeit usw.). Um diese Verschiedenartigkeit zu beseitigen, gibt es zwei Wege: entweder man ordnet bei jedem Test die vorliegenden Arbeiten nach den ursprünglichen Wertziffern; in dieser Ordnung erhält jeder Schüler für seine Arbeit einen Rangplatz, und diese Rangnummern sind nun bei den verschiedenen Tests vergleichbar. Oder man veranlaßt die Beurteiler, die Wertziffern des Tests in die üblichen Schulprädikate 1—5 (wo 1 = sehr gut, 5 = ungenügend) umzusetzen; und wiederum sind nun alle Tests auf einen gemeinsamen Generalnenner gebracht. Nunmehr ist es auch möglich, sämtliche Testergebnisse des einzelnen Prüflings zu einem Gesamtwert zu vereinigen und ihm gleichsam eine „Testzensur" oder einen Rangplatz für die Gesamtprüfung zu verleihen. Es ist dies der letzte Punkt, bis zu dem die Massenprüfung und ihre Verarbeitung zu führen vermag. Was nun weiter zu geschehen hat, wird in dem Abschnitt „Die Entscheidung" behandelt.

Zuvor aber müssen wir uns mit den andersartigen Prüfungsbedingungen beschäftigen, die bei Einzelprüfungen vorliegen. Überall wo die Zahl der Anwärter nur gering ist, wird man — zum mindesten neben einer schriftlichen Gruppenprüfung — auch die individuelle Prüfung jedes einzelnen durchzuführen suchen; insbesondere wird eine solche überall dort nötig sein, wo es sich nicht mehr nur allgemein um höheren oder geringeren Begabungs-

grad, sondern um besondere Begabungsrichtung und individuelle Berufseignung handelt. Endlich wird die Einzelprüfung als Nachkontrolle dort nötig werden, wo die Massenprüfung und die anderen Kriterien zu eindeutiger Entscheidung nicht ausreichen.

Die Einzelprüfung vermag naturgemäß sehr viel tiefer in die Eigenart der Veranlagung einzudringen. Zunächst ist die Zahl der verfügbaren Tests bedeutend größer, denn die in der Massenprüfung unmöglichen Formen der mündlichen Befragung und der stummen Handlungstests (z. B. Ordnungstests, Zusammensetztests, Ausführen von Aufträgen, Prüfung der praktischen Intelligenz, konstruktive Aufgaben usw.) sind nun anwendbar. Sodann besteht die Möglichkeit des Individualisierens: wird eine Aufgabe zunächst nicht gelöst, so kann man bestimmte Hilfen geben; wird sie spielend leicht gelöst, so kann man zu schwereren übergehen; an bestimmte Arten von Antworten und Lösungen kann man weitere Fragen knüpfen, die in anderen Fällen nicht möglich sind usw. Vor allem aber hat man für die psychologische Deutung der gegebenen Leistung ganz andere Grundlagen als im schriftlichen Versuch. Denn nunmehr kann man ja den Prüfling *während* der Arbeit selbst beobachten und viele charakteristische Züge: Planmäßigkeit oder Sprunghaftigkeit, Zögern oder Stetigkeit, das Selbstfinden von Tricks und Hilfen, Nachprüfung und Selbstkorrektur der eigenen Lösungsversuche usw. feststellen; außerdem kann man sofort durch geschickte Befragung Aufklärung gewinnen, was mit bestimmten Antworten und Lösungen gemeint sei und welche Gedankengänge zu ihnen hingeführt haben. Hierbei darf man freilich eines nicht vergessen: je stärker eine Prüfung individualisiert, um so weniger vergleichbar sind ihre Ergebnisse mit anderen Prüfungen. Diese Vergleichbarkeit aber dürfen wir nie aus dem Auge lassen — selbst dann nicht, wenn wir es überhaupt nur mit einem einzigen Prüfling zu tun haben. Denn wenn dieser Prüfling z. B. vom Psychologen einen Rat erhalten will, ob er auf Grund seiner Sonderbegabung in die Vorbildung zu einem technischen Beruf eintreten soll, so müssen wir wissen, ob seine technisch-mathematische Begabung die des *Durchschnitts* seines Alters wesentlich übertrifft; mit anderen Worten: wir müssen das Verhalten von Begabten und Unbegabten gegenüber den gleichen Prüfungsansprüchen kennen. Noch deutlicher wird dies dort, wo aus einer größeren Zahl von Bewerbern — etwa solchen, die in die Werkschule eines technischen Betriebes eintreten wollen — die Befähigtsten ausgelesen werden sollen; hier

wird zwar jeder einzelne geprüft, aber doch in steter Übereinstimmung mit den Prüfungen der Mitbewerber.

Aus diesen Gründen ist es methodisch empfehlenswert, die individualisierenden Abwandlungen der Prüfung nicht wahllos und je nach Bedürfnis vorzunehmen, sondern von den gleichförmigen Maßnahmen zu trennen. Man prüfe zunächst den einzelnen Prüfling möglichst exakt nach dem allgemeinen Prüfungsschema, so daß die hierüber vorliegenden Ergebnisse nach Grad und Art ohne weiteres mit denen der anderen verglichen werden können; erst wenn dieser Teil der Prüfung abgeschlossen ist, gehe man zu den individualisierenden Maßnahmen über, stelle besondere Fragen, oder lasse eine zuerst nicht gelöste Aufgabe noch einmal unter Darbietung von Hilfen vornehmen. So kann das erste mehr schematische (und vergleichbare) Ergebnis feiner herausmodelliert und zu einem individuellen Begabungsbild des Prüflings ausgebaut werden. Die Ausgestaltung der Einzelprüfung zu einem derart „kombinierenden" Verfahren wird im wesentlichen erst von der Zukunft zu erwarten sein.

3. Die Entscheidung.

„Die Entscheidung muß auf einer Gesamtberücksichtigung aller gewonnenen Kriterien beruhen" (s. S. 252).

An solchen Kriterien stehen folgende zur Verfügung: 1. Der abgebende Lehrer benennt den Schüler als etwaigen Anwärter für die Auslese, mit oder ohne besondere Empfehlung. 2. Die Schulzeugnisse des Schülers liegen vor. 3. Der Lehrer reicht ein Psychogramm des Schülers ein. 4. In der aufnehmenden Schule findet eine Aufnahmeprüfung in üblicher Form („Fertigkeitsprüfung") statt. 5. Es findet eine Auslesеprüfung in neuer Form („Fähigkeitsprüfung") statt, und zwar als Massen- oder Einzelprüfung.

Daß die ersten drei Hilfsmittel, die von der abgebenden Schule stammen, allein nicht ausreichen, um die Auslese zu bewerkstelligen, ist schon ausführlich an anderer Stelle (S. 253ff.) begründet worden. Es sei noch der sehr einfache aber durchschlagende Grund angeführt, daß ja die Anzahl der von den abgebenden Lehrern vorgeschlagenen Schüler und die Anzahl der in den Aufnahmeschulen verfügbaren Plätze kaum je übereinstimmen werden. Nur dann, wenn die erste Ziffer kleiner ist, würde sich eine besondere Auslesemaßnahme erübrigen; in allen anderen Fällen aber muß dem Vorschlag der abgebenden Lehrer noch eine weitere Siebung folgen, die nur in einer Prüfung bestehen kann. Die Gerechtigkeit erfordert, daß diese Prüfung für alle Beteiligten unter gleichen Bedingungen steht; die äußeren Umstände bringen es mit sich, daß die Prüfung oft nur als Massenprüfung durchführbar ist.

Aber auch diese Prüfung darf nicht für sich allein die Entscheidung herbeiführen. Dies gilt insbesondere von der psychologischen Testprüfung, die stets nur ergänzende, nicht aber allein ausschlaggebende Bedeutung gewinnen darf. Hierin weicht also unser Standpunkt von jenem ab, den Moede und Piorkowski (*206*) bei der Berliner Auslese vertreten haben. Wir haben oben die Mängel besprochen, die der Massenprüfung als solcher anhaften; wer diese Mängel anerkennt und richtig einschätzt, kann nicht mehr davon überzeugt sein, durch sie allein „empirische und unangreifbare Fähigkeitsreihen" zu erlangen (Moede-Piorkowski, S. 170). Empirisch ist die Testrangordnung wohl, indem sie jene Erfahrungen vereinigt, die eben auf dem Prüfungswege zu erzielen sind; aber eine andere Gruppe möglicher Erfahrungen ist hierbei vernachlässigt, nämlich jene, die nur auf dem Wege der Beobachtung gemacht werden können. Der so wichtige Unterschied zwischen reaktiven und spontanen Äußerungen von Fähigkeiten ist von M. und P. übersehen worden. Die experimentelle Prüfung hat es nur mit Reaktionen zu tun; jene Seiten der Begabung, die im natürlichen Ablauf des Lebens und im freien Spiel spontaner Interessenbetätigungen hervortreten, sind überhaupt nicht berücksichtigt. Darum ist — selbst wenn man eine fehlerlose Durchführung des experimentellen Verfahrens voraussetzt — jene Rangreihe durchaus noch nicht „unangreifbar"; und die Zuversicht, daß die Leistungen in der Begabungsschule mit dieser rein experimentell begründeten Diagnose übereinstimmen müßten, darf nicht ohne weiteres als gerechtfertigt gelten.

Ein anderes Argument gegen die Überwertung der Massentestprüfung liegt in dem mechanischen Charakter der Berechnung. Gewiß erscheint es sehr exakt, daß jeder Schüler seine Testzensur und nach dieser seine Rangnummer erhält; aber diese Exaktheit verliert doch so manches von ihrer Bedeutung, wenn man bedenkt, daß die Ziffer durch die Vereinigung vieler Teilziffern zustande gekommen ist, die nicht immer gleiche Wichtigkeit besitzen. Jedenfalls ist diese Exaktheit nicht so scharf, daß geringfügigen Unterschieden in diesen Resultantenwerten schon ein psychologisch bedeutsamer Unterscheidungswert zugemessen werden könnte; und wenn trotzdem so geringfügige Unterschiede über Lebensschicksale entscheiden können, so darf man nur noch von einem, mit dem Schein der Exaktheit umkleideten Würfelspiel sprechen. „Wir zählten von der resultierenden endgültigen mitt-

leren Rangreihe einfach die ersten 30 oder 60 ab und empfahlen sie zur Aufnahme in die Begabtenschulen." So berichten Moede und Piorkowski. Ja, einfach ist das Verfahren gewiß! Die um den Rangplatz 60 herum befindlichen Kinder mögen sich in ihren Gesamtleistungen sehr nahestehen; aber die Abzählungsmechanik entscheidet, und Nr. 60 wird Gymnasiast, während Nr. 61 Volksschüler bleibt. Vielleicht hat jener 61. gerade nur in einem Test, der mit wirklicher geistiger Fähigkeit herzlich wenig zu tun hat (z. B. im Lernen sinnloser Silben), etwas weniger geleistet als der 60. — solche psychologischen Erwägungen sind gegenüber der Allmacht der Gesamtziffer völlig zum Schweigen verurteilt.

Je eher wir uns von diesen Übertreibungen in der Bewertung des Tests losmachen, um so eher können wir verlangen, daß seine Bedeutung als dienendes Glied in der Reihe der Kriterien anerkannt werde.

Unsere ganze bisherige Betrachtung führt zu der zwingenden Forderung, daß eine neutrale Entscheidungsstelle geschaffen werde, an der alle von Lehrern und Psychologen erzielten Befunde zusammenlaufen, um für jedes Kind miteinander verglichen und gegeneinander abgewogen zu werden. Praktische Pädagogen und psychologisch geschulte Kräfte müssen an dieser Stelle zusammenwirken. Die Arbeit ist groß, aber nicht unüberwindlich. Denn für die meisten der Prüflinge werden die verschiedenen Kriterien gut zusammenstimmen: bei Kindern mit vorzüglichen Lehrerempfehlungen und Zeugnissen, günstig klingenden Beobachtungsbogen und guten Rangplätzen in der Begabungsprüfung ist die Aufnahme ohne weiteres gegeben; andere, bei denen die verschiedenen Kriterien gleichmäßig ungünstig liegen, sind unzweifelhaft abzuweisen. Aber zwischen diesen Gruppen liegt nun das breite Mittelgebiet der zweifelhaften Fälle; und gerade in diesem Gebiet, durch welches die bloße Testprüfung einen mechanisch errechneten Schnitt legen würde, bewährt sich das abwägende Verfahren aufs beste. Ein an sich auf der Grenze stehender Testausfall kann durch den Beobachtungsbogen nach der einen oder anderen Seite geklärt werden; es kann auch so manches Kind, das nach dem Testausfall allein abzuweisen wäre, im Beobachtungsbogen oder in den Zeugnissen Kompensationen liefern, die seine Aufnahme rechtfertigen; auch die Aufnahme auf Probe oder die Verweisung des Falles an eine persönliche Nachprüfung des Kindes nach dem individuellen Testverfahren kann ausgesprochen werden.

Mir scheint, daß auf solche Weise das Höchstmaß an Zuverlässigkeit gewonnen wird, das bei pädagogisch-psychologischen Entscheidungen überhaupt und bei dem Stand unserer Methoden zur Zeit erreichbar ist.

Eine solche neutrale Entscheidungsstelle muß aber auch da bestehen, wo die Auslese im allgemeinen ohne testpsychologische Prüfungen vorgenommen wird. Denn Zweifelfälle muß es bei jeder Auslese geben, und zum mindesten bei diesen muß mit aller Entschiedenheit die Heranziehung psychologischer Gesichtspunkte zur Entscheidung gefordert werden.

Setzen wir etwa den Fall, in irgend einer Stadt begnüge man sich bei der Überführung von Volksschülern in die Unterstufe der höheren Schulen mit den Benennungen und Beobachtungen der abgebenden Lehrer und der dann folgenden gewöhnlichen (d. h. auf Kenntnisse gerichteten) Aufnahmeprüfung. Dann werden stets Fälle vorkommen, in denen sich Eltern mit der Ablehnung ihrer Kinder nicht zufrieden geben, oder andere, in denen die aufnehmenden Lehrer die Empfehlung des abgebenden Lehrers nicht anerkennen. Für solche vermutlich nicht allzu seltenen Fälle muß eine unparteiische Prüfungsinstanz vorhanden sein, die nun ihrerseits versuchen muß, ihre endgültige Entscheidung auf eine individualisierende Befähigungsfeststellung zu gründen.

Einfacher liegt die Organisation der Entscheidung dort, wo es sich nicht um große Massen von Kindern handelt, sondern wo eine kleinere Zahl zur engeren Wahl steht: z. B. bei der Aufnahme in eine Fachschule. Hier kann ja jeder Kandidat individuell behandelt werden; und da wird es vor allem darauf ankommen, die Aufnahmeprüfung selbst psychologisch zu vertiefen. Man kann hier entweder neben die — ja nicht zu entbehrenden — Kenntnisprüfungen Fähigkeitsprüfungen setzen, um zugleich ein Bild der erworbenen Fertigkeiten und eines der angeborenen Veranlagungen zu gewinnen; oder man kann geradezu Gesichtspunkte der Begabungsfeststellung in die Kenntnisprüfung selbst hineinverarbeiten (indem man z. B. eine Rechenarbeit dazu benutzt, die Kritikfähigkeit gegenüber widersinnigen Aufgaben zu prüfen). Vielleicht liegt in dieser Richtung — also in der Verbindung von Kenntnis- und Fähigkeitsprüfung und damit in einer Reform des Prüfungswesens — die Hauptaufgabe, die der Begabungspsychologie auf praktischem Gebiet bevorsteht. Denn daß die schroffe Gegenüberstellung

pädagogischer Prüfung und psychologischer Prüfung nur für die erste Phase dieser neuen Entwicklung möglich war, erscheint mir ganz sicher; wir müssen mit der Idee eines einheitlichen „pädagogisch-psychologischen" Ausleseverfahrens rechnen.

Die neutrale Prüfungsstelle, von der oben die Rede war, wird übrigens ihre Arbeit nicht auf die letzte Entscheidung über vorgenommene Auslesen beschränken können. Sie wird ständig auch die Vorbereitung der Auslese zu verwalten, Tests auszuarbeiten, deren Bewertung zu überwachen, Beobachtungspläne zu entwerfen und ihre Ausfüllung zu studieren, endlich auch die ausgelesenen Kinder in ihrer weiteren Entwicklung zu verfolgen haben. Sie wird sich, namentlich in größeren Städten, zu einem ständigen Ausleseamt der Schulbehörden entwickeln, kurz, zu einem „schulpsychologischen Amt", dessen Notwendigkeit schon vor Jahren von mir entwickelt worden ist[1]).

4. Erprobung und „bewegliche Front".

Mit der Zuweisung eines Kindes zu einer Begabungsklasse oder höheren Schule ist nun aber das Ausleseverfahren für diesen Schüler nicht endgültig abgeschlossen. Denn trotz aller Vorsichtsmaßregeln sind die Auslesemethoden unsicher; außerdem kann die Entwicklung eines Kindes stets eine unerwartete Wendung zum Besseren oder Schlechteren hin nehmen. Wir müssen deshalb darauf gefaßt sein, die ursprüngliche Zuweisung revidieren und auch unsere Auslesemethoden auf Grund der weiterhin gemachten Erfahrungen verändern zu müssen.

Es gibt eine Reihe von Pädagogen, welche sich gegen jedes eigentliche Ausleseverfahren wenden und alles der „Erprobung" anheimstellen wollen. Man solle alle Schüler, die für die höheren Züge gemeldet werden, unterschiedslos aufnehmen, aber das erste Jahr als Probezeit betrachten; wer sich dann als ungeeignet erwiesen habe, werde wieder an die Ausgangsschule zurückgegeben.

[1]) Zuerst 1911 in einem Kongreßvortrag (*12a*), dann ausführlicher unter Aufklärung gewisser inzwischen hervorgetretener Mißverständnisse in „Jugendkunde als Kulturforderung" (*14*). Aus dem Meinungsaustausch, der sich im wesentlichen um Begriff und Aufgabe des „Schulpsychologen" drehte, seien hier genannt: die zustimmenden Ausführungen von Mann (*202*), Hylla (*192*), v. d. Wyenbergh (*243*), die polemischen von Kosog (*195, 196*), Weigl (*239*), Ehlers (*187*), und meine eigene (*229*). Über eine dem gleichen Thema geltende Aussprache auf der erzieh.-wiss. Hauptstelle des Deutschen Lehrervereins berichtet die Päd. Ztg. (*188*).

Diesem Vorschlag, der nur aus einer blinden Gegnerschaft gegen jedes Auslese- und Prüfungsverfahren verständlich ist, muß mit aller Entschiedenheit entgegengetreten werden. Um die Alleinherrschaft der schulmeisterlichen Empirie zu retten, sollen unsere Schulen zunächst mit möglicherweise Ungeeigneten überfüllt und soll ein Lebensjahr von zahlreichen Kindern unnütz vertan werden. Welche Sinnlosigkeit wäre es, in einer Großstadt, die etwa 2000 Schüler in den Aufnahmeklassen der höheren Schulen unterbringen könnte, nun plötzlich lauter neue Klassen nur deshalb zu schaffen, weil 3000 Schüler vorgeschlagen worden sind — und dann nach einem Jahr die neuen Klassen wieder aufzulösen und 1000 Schüler an die Volksschulen zurückzuverweisen! Hat man denn kein Verständnis dafür, welche Enttäuschung und Verbitterung diese nachträgliche Zurückweisung erzeugt und welchen schweren Stand diese Kinder dann in den Volksschulen haben werden? Eine Erprobung ist gewiß notwendig; aber sie ist ein notwendiges Übel, dessen Umfang möglichst verringert werden muß; und der tiefste Sinn des psychologischen Ausleseverfahrens besteht ja eben darin, rechtzeitige Sonderungen der Geeigneten und Ungeeigneten vor dem Eintritt in die Aufnahmeschule vorzunehmen, um die Zahl der Fehlentscheidungen und später notwendigen Rückgängigmachungen auf ein Mindestmaß zurückzuführen. Auch hier ist, wie überall, Prophylaxe besser als Therapie!

Zieht man, wie wir es fordern, bei der Auslese selbst alle verfügbaren Hilfsmittel heran, dann wird nur ein kleiner Prozentsatz von zweifelhaften Fällen übrig bleiben, deren Aufnahme ausdrücklich als „probeweise" bezeichnet werden muß; und nach einem Jahr werden unter diesen wie unter den anderen nur wenige Kinder zu finden sein, die sich nicht bewährt haben; diese müssen natürlich in einen anderen, ihren nun besser erkannten Anlagen entsprechenden Bildungsgang überführt werden. Bei den ersten noch unzulänglichen Auslesemethoden ist der Prozentsatz der Nichtbewährung noch ziemlich beträchtlich (wofür die Berliner Begabtenschulen bereits das Beispiel geben); je mehr sich das Ausleseverfahren verfeinert, um so geringer werden die Nötigungen zu nochmaligem Wechsel sein (was soeben durch die niedrigen Abschulungsziffern in Hamburg belegt wird); aber ganz werden sie nie aufhören, schon wegen der oben erwähnten unvorhersehbaren Schwankungen und Wendungen in der geistigen Entwicklung.

Gerade der Psychologe hat allen Anlaß zu betonen, daß von

„Sicherheit" bei psychologischen Prognosen und Diagnosen Jugendlicher niemals gesprochen werden kann. Die Entwicklungsgänge werdender Persönlichkeiten lassen sich nicht vorausbestimmen wie astronomische Gestirnbewegungen; und noch so fein ausgearbeitete Beobachtungs- und Prüfungsmethoden können niemals mehr als Wahrscheinlichkeiten bezüglich der künftigen Fähigkeits- und Interessenentwicklung geben.

Diese psychologische Einsicht führt dazu, an die pädagogische Reform eine ganz bestimmte Forderung zu stellen: niemals darf eine einzige Auslese auf einer bestimmten Altersstufe endgültig und unwiderruflich über das Schul- (und damit Lebens-)schicksal des Kindes entscheiden. Der Aufbau der gegliederten Einheitsschule muß vielmehr beweglich sein in doppelter Hinsicht. 1. Da sich mit steigendem Alter die speziellen Befähigungen und Interessen immer deutlicher bekunden, so sind wiederholte Auslesen mit immer engerer Zielbestimmung nötig. Die Gliederung der Einheitsschule wird eine stufenweise sein müssen in dem Sinne, daß zunächst (etwa nach vier Schuljahren) nur die Kinder von höherer Allgemeinintelligenz, die überhaupt für eine höhere Schulbahn in Betracht kommen, ausgelesen werden, und daß erst einige Jahre später die Trennung der verschiedenen Schulzweige (der humanistischen, realistischen, technischen, künstlerischen usw.) auf Grund der Sonderbegabung erfolgt. 2. Für diejenigen Schüler, bei denen sich erst nach dem Normalauslesetermin die Eignung für eine andersartige Schulgattung oder die Ungeeignetheit zum Verbleiben in der bisherigen herausstellt, muß stets die Übergangsmöglichkeit geschaffen werden; es sind dazu durch Sonderkurse und ähnliche Einrichtungen Querverbindungen zwischen den verschiedenen Zweigen der Einheitsschule herzustellen. Die weitere Durcharbeitung dieser Forderung muß der pädagogischen Organisation überlassen bleiben; der Psychologe muß sich mit der Aufstellung des Grundsatzes begnügen. Aber soviel ist klar, daß alle diese Sonderfälle wiederum ein besonderes Arbeitsgebiet jenes neutralen Ausleseamtes bilden werden, dessen Notwendigkeit oben erwiesen wurde.

Die Erprobung muß sich schließlich — und das geht wieder in erster Linie den Psychologen an — auf die Methoden der Auslese beziehen. An den Erfahrungen, die mit den ausgelesenen (und auch mit den nicht aufgenommenen) Schülern weiterhin gemacht

werden, ist stets zu studieren, ob die Seelenbilder der „Beobachtungsbogen" sich bestätigen, ob die Rangordnungen der Leistungen mit den Testergebnissen in Einklang stehen, welche Tests vor allen Dingen durch die Praxis bestätigt, welche durch sie verleugnet werden. Diese letzte Frage ist von besonderer Wichtigkeit für die ganze Zukunft der psychologischen Auslese. Denn Beobachtungsbogen und Tests sind ja dazu bestimmt, prognostischen Wert zu erhalten; und nur dann, wenn sie sich so bewähren, daß man ihnen zum mindesten eine prognostische Wahrscheinlichkeit höheren Grades zusprechen kann, haben wir das Recht, ihre endgültige Einführung zu fordern. Diese methodischen Erprobungen werden sich natürlich durch mehrere Jahre hinziehen; und so ist denn auch von diesem Gesichtspunkt aus der nur vorläufige Charakter der bisherigen psychologischen Auslesemethoden ausdrücklich zu betonen. Daraus ist natürlich keinerlei Beweisgrund gegen ihre Anwendung zu entnehmen; im Gegenteil: um überhaupt in absehbarer Zeit zu der so nötigen Sicherung des Ausleseverfahrens zu gelangen, muß schon jetzt in möglichst weitem Umfange dafür Sorge getragen werden, daß Methoden verschiedener Art angewandt und erprobt werden. Schon im gegenwärtigen Zustand der Unvollkommenheit und Vorläufigkeit sind sie doch besser als gar keine Methoden; sie sind zum mindesten bereits ein vielversprechender Anfang, um das Würfelspiel, das in jeder Auslese liegt, zu mildern, und den Aufstieg der Tüchtigen zu regeln[1]).

Noch weit über die eben geforderten Erprobungen hinaus gehen die Vorschläge, die M. Vaerting soeben macht, um „Wege zur Gewinnung neuer Maßstäbe für die Begabungsprüfungen" zu gewinnen (*238a*). Nicht ganz mit Unrecht bemängelt er an den bisherigen Methoden, daß man ihren Wert letzten Endes immer wieder an den Urteilen der Schule geeicht habe, sei es an den gleichzeitigen (im Beobachtungsbogen), sei es an den nachfolgenden (der Bewährung in den Begabtenklassen). Vaerting ist von starkem Mißtrauen gegen die bestehende Schule erfüllt bezüglich ihrer Fähigkeit, die wahren Begabungen zu erfassen; ja sie sei sogar

[1]) Mit Recht wendet sich auch Brahn (*184*) in seinen „Besinnlichen Betrachtungen zur Begabungsprüfung" gegen diejenigen, die aus den bisherigen Mängeln dieses Verfahrens seine grundsätzliche Ablehnung folgern wollen. Er schließt, ähnlich wie ich, mit einem Appell an die Behörden, eine Auslesekonferenz für ganz Deutschland einzuberufen.

geeignet, etwa vorhandene Begabungen zu unterdrücken. Deshalb solle man auch die experimentellen Begabungsprüfungen nicht zu sehr den Schulforderungen anpassen, und man müsse vor allem als ihren wahren Erprobungsmaßstab nicht die Schulbewährung, sondern die Lebensbewährung ansehen. Darum verlangt er eine großzügige, erst in Jahrzehnten durchzuführende Organisation der Begabungsforschung. Einerseits sollen hervorragende Persönlichkeiten, die ihre Begabung praktisch erwiesen haben, psychographiert und geprüft werden — möglichst mit analogen Methoden wie die Schüler —; nur so könne festgestellt werden, welche Prüfmittel für echte Begabung wahrhaft symptomatisch seien und welche nicht. Ferner sollen alle Personen, die einer Auslese- oder Eignungsprüfung unterzogen worden waren, Jahre und Jahrzehnte hindurch in ihrem Berufs- und Fähigkeitsschicksal verfolgt werden; erst hierdurch werde der prognostische Wert oder Unwert der damaligen Entscheidung festgestellt. Und zwar gelte diese Forderung nicht nur für die Aufgenommenen, sondern auch für die auf Grund der Prüfung Abgelehnten; mit Recht tadelt Vaerting, daß diese Nachprüfung der negativen Fälle bisher viel zu leicht genommen worden sei. V. fordert dann weiter eine besondere Eignungsprüfung der Lehrer und Psychologen, die bei dem so verantwortlichen Ausleseverfahren beteiligt sind.

Viele dieser Vorschläge klingen heute noch utopisch, zumal wenn man die vorhandenen kärglichen Mittel und spärlichen Arbeitskräfte der wissenschaftlichen Begabungsforschung mit dem Kraft- und Personalaufwand vergleicht, der für die Durchführung solcher Arbeiten nötig wäre. Wie viel oder wie wenig von diesen Forderungen sich mit der Zeit durch den sachlichen Zwang der Verhältnisse durchsetzen wird, läßt sich jetzt noch nicht sagen. Im Augenblick müssen wir uns begnügen, aus dem Wust der Probleme diejenigen zunächst herauszugreifen, deren Lösung die dringlichste Forderung des Tages darstellt und mit den verfügbaren Hilfsmitteln und -kräften möglich ist. In diesem Sinne sind meine eigenen Vorschläge zu verstehen.

Der Größe der Aufgabe aber und der Unvollkommenheit ihres gegenwärtigen Lösungsstandes würde es entsprechen, wenn die gründliche Durchprüfung des psychologisch-pädagogischen Ausleseverfahrens und seine allmähliche Überleitung in die Praxis einem Reichsausschuß übertragen würde, der aus Pädagogen, Psychologen, Ärzten, Sozialpolitikern und Berufsvertretern bestehen müßte.

XIII. Kapitel.

Übersicht über die bisherigen Schülerauslesen.

Die grundsätzlichen Betrachtungen des vorigen Abschnitts ermöglichen es, uns bei der Besprechung der bisher in der Öffentlichkeit bekannt gewordenen psychologisch-pädagogischen Schülerauslesen kürzer zu fassen. Daß diese nur erste Versuche auf bisher unbetretenen Pfaden darstellen und deshalb nicht frei von manchen Mängeln sind, wird von allen Beteiligten mit größerer oder geringerer Selbstkritik eingestanden. Dennoch werden die Erfahrungen, die dort gemacht worden sind, allen künftigen Arbeiten auf diesem schwierigen Gebiet von Nutzen sein können.

Die bisherigen im deutschen Sprachgebiet vorgenommenen Auslesen[1]) erstrecken sich auf zwei große Altersgruppen: auf Kinder

---

[1]) Ob im Ausland während der letzten Jahre ebenfalls Schülerauslesen mit Hilfe psychologischer Methoden vorgenommen worden sind, entzieht sich unserer Kenntnis. Nur zwei schon vor dem Kriege veröffentlichte Versuche aus Amerika und England seien anmerkungsweise erwähnt.

Martha Adler (*180*) prüfte zwei Klassen einer Neuyorker Volksschule. Zunächst wurden 70 Knaben der untersten Klasse, die $^1/_2$ Jahr in der Schule waren, nach den Ergebnissen der Binet-Prüfung in zwei Abteilungen zu je 35 geteilt, welche getrennt unterrichtet wurden. In Abteilung A konnte in schnellerem Tempo und mit weit besserem Erfolg unterrichtet werden als in B. Von den 35 A-Schülern gingen 5 im Laufe des nächsten Halbjahres ab; von den übrigen 30 waren 18 am Schluß des ersten Schuljahres reif für die zweite Hälfte des zweiten Jahreskurses; sie hatten also 6 Monate gewonnen. Aber auch bei den Zöglingen der Abteilung B hatte die größere Ähnlichkeit des geistigen Standes die Unterrichtsarbeit beträchtlich erleichtert.

Ferner wurden 89 Knaben des vierten Schuljahres geprüft, sowohl mit den Binet-Tests für 10jährige, wie mit einigen im Massenversuch anwendbaren Tests. Die 36 Schüler mit den besten Prüfungsergebnissen wurden zur Abteilung A vereinigt, in welcher außer dem normalen Pensum des vierten Schuljahres noch die Hälfte des fünften Jahrespensums bewältigt werden konnte. 22 von diesen konnten am Schluß des Jahres in die obere Hälfte des fünften Kurses überführt werden, sie hatten also wiederum ein halbes Jahr erspart. In Abteilung B war das normale vierte Jahrespensum erledigt worden.

St. Wyatt (*242*) stellte in Manchester an 10—13jährigen Schülern Intelligenzprüfungen mit verschiedenen Tests an. U. a. prüfte er 7 zur Aufnahme in eine höhere Schule bestimmte Schüler mit dem Analogietest und dem Lückenergänzungstest und konnte daraufhin bei zweien die Ablehnung, bei den 5 anderen die Zuweisung zu bestimmten Klassen empfehlen. Seine Wertung deckte sich nicht nur mit dem Ergebnis der eigentlichen Aufnahmeprüfung, sondern auch mit den späteren Schulerfolgen der geprüften Kinder.

im Alter von 9 und 10 Jahren und auf Jugendliche im Alter von 12 bis zu 15 Jahren. Unser Bericht beginnt mit den Jüngsten und schreitet zu den Älteren fort.

Die erste Auslese hat dort einzusetzen, wo sich aus der für alle Kinder gemeinsamen Grundschule der gegliederte Oberbau abzuzweigen beginnt. Als Zeitpunkt für einen solchen ersten Einschnitt im Schulsystem hatte sich schon vor Einführung der Einheitsschule der Abschluß des dritten Schuljahres (Übergang von der Vorschule zur höheren Schule) oder das Ende des vierten Schuljahres (Übergang von der Volksschule zur höheren Schule) eingebürgert; und es hat den Anschein, als ob in Zukunft der Abschluß des vierten Schuljahres allgemein einen solchen ersten Auslesetermin bilden wird[1]). Wie schon oben erwähnt, wird es für dieses frühe Alter noch nicht möglich sein, besondere spezifische Begabungsrichtungen bei den Kindern zu diagnostizieren; man wird sich im wesentlichen auf die Feststellung der allgemeinen Intelligenz beschränken müssen; höchstens wird man dort, wo es sich um die Auslese für Fremdsprachschulen handelt, die sprachliche Seite der Intelligenz besonders berücksichtigen können.

1. Auslese 9jähriger Kinder (Dr. Lipmann, Berlin).

Mit 9jährigen Mädchen hat es eine dem Umfange nach sehr begrenzte Auslese in Berlin zu tun, über welche Lipmann berichtet (*198*). Es handelte sich um die Aufnahme von Volksschülerinnen in ein städtisches Lyzeum; die Frage war, ob die Kinder, welche vier Jahresstufen der Volksschule bereits hinter sich haben, der bisherigen Übung entsprechend sämtlich erst in das vierte Schuljahr der höheren Schule eintreten müssen — was einen Zeitverlust von einem Jahr für sie bedingen würde — oder ob die Befähigteren unter ihnen bereits für die Klasse reif sind, die um ein halbes Jahr oder gar um ein ganzes Jahr höher liegt. Im letzten Falle würden die Kinder der Volksschule und der höheren Schule nach gleich langer Vorbildung auf gleicher Stufe stehen. Daß dann gewisse Abstände in Kenntnissen und Fertigkeiten vorhanden wären, ist freilich unbestreitbar; diese wären aber leicht nachzuholen, wofern der Stand der geistigen Reife der gleiche ist. Nur dieser sollte geprüft werden. Zur Vorbereitung der

[1]) Wenigstens für die Auslese zu höheren Schulformen hin. Die Auslese der Unternormalen zur Hilfsschule hin kann natürlich schon früher stattfinden.

Prüfung hat Lipmann vor allem die Leistungsfähigkeit der drei in Betracht kommenden Aufnahmeklassen für die anzuwendenden Tests geeicht, um danach die richtige Einordnung der aufzunehmenden Schülerinnen zu bestimmen.

L. wählte folgende Tests: 1. Bindewortergänzung; 2. Lehre aus einer Fabel ziehen; 3. Behalten von Einzelheiten einer Geschichte; 4. Abzeichnen geometrischer Figuren aus dem Gedächtnis; 5. Verständnis einer Situation; 6. Verstandesfragen. Bewertet wurde der Schwierigkeitsgrad jeder Lösung; hieraus wurden dann für jedes Kind „Intelligenzzensuren" abgeleitet. (Über diese Berechnungsmethoden vgl. Kap. V, 5c dieses Buches.)

Nunmehr konnte für jede der drei Aufnahmeklassen[1]) die Verteilung und Rangordnung der Intelligenzzensuren festgestellt werden; es zeigte sich ein durchgehender und deutlicher Unterschied der drei Klassen. Das nach der Intelligenzzensur schlechteste Viertel der Kinder wurde für jede Klasse abgegrenzt und hiernach auch die Grenze für die Aufzunehmenden gezogen. Diese Grenze lag z. B. für die oberste der drei Klassen bei der Intelligenzzensur 49; aufzunehmende Kinder, deren Testleistung diesen Wert nicht erreichten, die also dem schlechtesten Viertel der Klasse gleichstanden, galten daher als nicht reif für diese Klasse. Bei der zweiten der Aufnahmeklassen war der Grenzwert 27; wer auch diesen nicht erreichte, gehörte in die unterste Aufnahmeklasse.

Über die Bewährung dieser Methode sind noch keine Mitteilungen erfolgt.

2. Auslese 9—11jähriger Kinder (Dr. Bobertag).

Im Frühjahr 1919 veranstaltete Dr. Bobertag (*183*a) am Arndt-Gymnasium in Dahlem bei Berlin eine Testprüfung zu dem Zweck, die übliche Aufnahmeprüfung für die unteren Klassen zu ergänzen. Die angewandten Tests wurden kurz vor dem Aufnahmetermin in den drei Klassen, für welche die Aufnahmen stattfinden sollten, durchgeprobt, um genaue Vergleichsmaßstäbe für die neu Eintretenden zu bieten.

Folgende Aufgaben wurden gestellt:

1. Gedächtnis für sinnvoll zusammenhängende Begriffsgruppen (z. B. Frühling, Sonne, Knospe).

2. Finden von Analogien (z. B. Schuster — Stiefel; Schneider — ?).

---

[1]) Geprüft wurden im September 1917 diejenigen Klassen, die Ostern 1918 die neu einzuschulenden Volksschülerinnen aufnehmen sollten; es waren die Klassen VIII M, VIII O, IX M mit dem durchschnittlichen Alter von 9; 5—8; 10—8; 3.

3. Ordnen von Figuren auf Grund eines selbst zu findenden Prinzips (Größe, Helligkeit).

4. Dreiwort-Methode. Aus jeder Wortgruppe sind möglichst viel verschiedene Sätze zu bilden.

5. Zu einem gegebenen Wort (z. B. Soldat) möglichst viel damit zusammenhängende Begriffe zu finden.

6. Finden von Oberbegriffen zu gegebenen Worten.

Jede Leistung erhielt eine Punktziffer, jeder Schüler auf Grund seiner sämtlichen Punktziffern einen Rangplatz. Die so entstehende Rangordnung hatte mit der Intelligenzschätzung durch den Lehrer eine hohe Korrelation, so daß die Testserie mit Recht als geeignetes Symptom der Intelligenz für die Aufnahmeprüfung verwertet werden konnte. Sie wurde deshalb neben der üblichen Kenntnisprüfung verwandt und diente dazu, diese zu kontrollieren oder unter Umständen zu korrigieren. Namentlich konnte bei gutem Ausfall der Testprüfung die Aufnahme trotz geringerer Kenntnisse ausgesprochen werden.

3. Auslese 10jähriger Kinder (Hamburg).

Die weitaus umfänglichsten Erfahrungen liegen bisher bei 10jährigen Kindern vor, also bei jener Altersstufe, die vermutlich auch in Zukunft Gegenstand der ersten Auslese sein wird. Die Erfahrungen sind vor allem in Hamburg gesammelt worden (*218*).

Wir geben zunächst einen kurzen Tatsachenbericht über das Hamburger Verfahren, um dann über Ergebnisse, Schwierigkeiten und Bewährung einige Bemerkungen anzuknüpfen.

In Hamburg wurde im Jahre 1918 der neue Typ einer gehobenen Volksschule geschaffen, der nach vier Schuljahren von der gewöhnlichen Volksschule abzweigt und in fünf weiteren Schuljahren (statt der sonst üblichen vier) die Schüler zu einer vertieften Bildung führen soll. Der neue Zug unterschied sich vom Normalzug vor allem durch die Einbeziehung von Fremdsprachen (Fremdsprachzug oder „F-Zug"). Für die Anfangsklassen dieser F-Züge waren nun aus sämtlichen Knaben- und Mädchenvolksschulen die Geeignetsten auszuwählen, in der Höhe von etwa 5%. Da die Zahl der Hamburger Volksschüler des vierten Schuljahres beiderlei Geschlechts im Jahre 1918 ungefähr 20000 betrug, so waren rund 1000 Kinder auszulesen.

Die Hamburger Oberschulbehörde hielt hierfür die Anwendung eines pädagogisch-psychologischen Ausleseverfahrens für notwendig

und setzte sich zu diesem Behufe mit dem psychologischen Laboratorium in Verbindung; im Zusammenarbeiten beider Instanzen wurde eine Organisation geschaffen, die Ostern 1918 zum ersten Male, Ostern 1919 zum zweiten Male in etwas veränderter Form in Funktion trat. Die grundsätzlichen Erörterungen des vorigen Abschnitts sind wesentlich den Erfahrungen — und zwar ebenso den positiven Erfolgen wie den Schwächen und Reibungen — entnommen, die bei dieser Auslesearbeit und bei der Vorbereitung künftiger Auslesen hervortraten.

Die Auslese für 1918 nahm folgenden Verlauf[1]).

In dem vorhergehenden Jahre arbeitete das psychologische Laboratorium sowohl einen psychologischen Beobachtungsbogen wie eine Reihe neuer Tests aus. Der Beobachtungsbogen wurde hauptsächlich von Martha Muchow (s. S. 261) entworfen und vorgeprüft: äußere Verhältnisse bewirkten, daß er erst wenige Wochen vor dem Auslesezeitpunkt den abgebenden Lehrern der vierten Volksschulstufen übergeben werden konnte. Diese erhielten die Weisung, aus jeder Knabenklasse bis zu 6, aus jeder Mädchenklasse bis zu 3 Kindern für die F-Züge vorzuschlagen, für diese Kinder Beobachtungsbogen auszufüllen und die ihnen besonders geeignet erscheinenden Kinder als solche zu kennzeichnen („empfohlene Kinder"). Bei dieser Vorauslese wurde aus früher erörterten Gründen (s. S. 255) die Zahl der erlaubten Nennungen wesentlich größer genommen als die Zahl der Aufnahmemöglichen betrug. Genannt wurden im ganzen 1355 Kinder (882 Knaben, 473 Mädchen). Von diesen sollten 14 Knaben- und 8 Mädchen-F-Klassen gebildet werden mit insgesamt 990 Kindern. Aus den Nennungen der Vorauslese waren also 73% endgültig auszulesen.

Die 1355 wurden einer Testprüfung unterworfen, die natürlich auf Individualuntersuchung verzichten mußte. Der äußere Grund, daß nur schriftliche Arbeiten möglich waren, und der innere Grund, daß für Sprachklassen ausgelesen werden sollte, war dafür bestimmend, vornehmlich sprachlich-logische Tests zu wählen. Jeder einzelne der anzuwendenden Tests sowie der Wortlaut der Anweisungen war von Mitgliedern des psychologischen Seminars vorbereitet und geeicht und hierbei vielfach geändert worden. Diese Vorprobung mußte in Vororten von Hamburg vorgenommen werden, um ein vorzeitiges Bekanntwerden auszuschließen. Gewöhnlich

[1]) Vgl. die ausführliche Darstellung: R. Peter und W. Stern (*218*).

wurde der zu eichende Test an normalen 10-, 11- und 12jährigen Kindern ausgeprobt: Tests, die hierbei steilen Altersfortschritt zeigten und bei 12jährigen Normalen ausreichende Leistungen aufwiesen, schienen zur Feststellung 10jähriger Hochbegabter geeignet. Es ergab sich schließlich folgende Testreihe:

a) Analytische Aufgaben.

Erklärung der 4 Begriffe Beute, Onkel, Mut, Neid.

Fabeltest: Aus zwei gegebenen Fabeln mußte die „Lehre" gezogen werden.

Kritiktest: Die in eine Geschichte eingestreuten Sinnwidrigkeiten mußten gefunden und kritisiert werden.

Merkfähigkeitstest: Lange, einmal dargebotene Sätze sollten ans dem Gedächtnis niedergeschrieben werden; entscheidend war, ob nur mechanisch Teile des Wortlautes oder judiziös der Sinn (eventuell in anderer Formulierung) wiedergegeben wurde.

b) Synthetische Aufgaben.

Bindewortergänzung: Der bekannte Ebbinghaussche Versuch in der von Lipmann vorgeschlagenen Abänderung: es sind im durchbrochenen Text lediglich die Bindewörter weggelassen, also die Träger des Zusammenhangs der einzelnen Sätze und Satzteile.

Dreiwortmethode: Aus drei gegebenen Wörtern einen sinnvollen Satz zu bilden. Die Beispiele werden den Piorkowskischen Versuchen entnommen.

Begriffe ordnen: Eine Reihe von Begriffen, die Bestandteile eines kausalen Zusammenhanges bilden, aber wahllos durcheinander gemischt waren, sind in die richtige Ordnung zu bringen.

Aufsatztest: Zwei Bilder, die zwei Phasen einer zusammenhängenden Begebenheit darstellten, wurden vorgelegt; die Begebenheit mußte erkannt und in einem Aufsatz dargestellt werden.

Um die gleichzeitige und völlig gleichmäßige Prüfung mit dieser Testreihe zu ermöglichen, wurden die 1355 Prüflinge in 61 Gruppen von je 20 bis 25 Kindern geteilt; jede Gruppe wurde einem Prüfer übergeben. Als Prüfer hatten sich Lehrer und Lehrerinnen zur Verfügung gestellt, die als Teilnehmer der psychologischen Vorlesungen und Übungen auf die Probleme eingestellt waren; sie wurden vorher mündlich und schriftlich über jede Maßnahme der Prüfung auf das genaueste unterrichtet. Die Prüfung währte vier Stunden, die auf zwei Vormittage verteilt waren.

Sofort nach Beendigung der Prüfung wurde für jeden der acht Tests eine Arbeitsgruppe gebildet, die unter Leitung eines Mitarbeiters des psychologischen Seminars an die Bewertung heranging. In jeder dieser Gruppen wurden die Bewertungsgrundsätze für den übernommenen Test festgestellt; um Gleichmäßigkeit zu erzielen, wurden überall die Bewertungsergebnisse in die be-

kannten Schulprädikate 1—5 zusammengefaßt. Binnen einer Woche waren von jeder Gruppe die 1355 Blätter ihres Tests korrigiert und bewertet zurückzuliefern. Die Zusammenstellung dieser Ziffern ergab nun für jeden Prüfling acht Prüfungsprädikate, die zu einem Gesamtprädikat zusammengefaßt werden konnten. (Hierbei wurden die Prädikate für die wichtigeren Tests: Aufsatz, Kritiktest und Bindewortergänzung, doppelt gezählt.) Die denkbar beste Leistung würde hiernach die Punktzahl 11 (in jedem Test Prädikat 1), die denkbar schlechteste die Punktzahl 55 (in jedem der Tests das Prädikat 5) erhalten. Innerhalb dieser Zahlgrenzen bewegen sich alle Gesamtprädikate, die nun in eine Häufigkeitstabelle gebracht wurden. Es ergab sich eine ziemlich symmetrische Verteilungskurve mit dem Häufigkeitsgipfel bei 34 Punkten, d. h. bei dem Prädikat 3; auch die Durchschnittsberechnung aus allen Gesamtprädikaten ergab den gleichen Wert. Wären die Testprüfungen zum alleinigen Maßstab erhoben worden, so würde die Punktzahl 38 als Grenzscheide gelten müssen; denn die Zahl derjenigen Kinder, die geringere Punktziffern als 38, d. h. bessere Prüfungsleistungen aufzuweisen hatten — 73% aller Prüflinge —, entsprach der Zahl der aufzunehmenden Kinder. Dies mechanische Abgrenzungsverfahren wurde aber aus den oben (S. 272/3) ausführlich erörterten Gründen nicht gewählt.

Vielmehr wurde die endgültige Entscheidung einem Ausschuß (von vier Pädagogen und einem Psychologen) übertragen, der jeden einzelnen der 1355 Fälle gesondert behandelte und hierbei alle verfügbaren Beurteilungsunterlagen heranzog. Zu diesem Zweck waren auch die ausgefüllten Beobachtungsbogen von einer besonderen Arbeitsgruppe durchgearbeitet worden; diese hatte kurze Auszüge und besonders bemerkenswerte Einzelheiten in die Schülerliste eingetragen. Außerdem standen dem Aufnahmeausschuß die Schulbogen zur Verfügung, in denen sämtliche bisherigen Schulzeugnisse übersichtlich enthalten waren; freilich zeigten diese keine große Differenzierung, da ja durchweg gute Schüler zur Beurteilung standen; die Zeugnisprädikate schwankten nur zwischen 1 und 2. Endlich wurde die besondere Empfehlung beachtet, die die Lehrer manchen Kindern durch Unterstreichung der Namen hatten zuteil werden lassen. Diese Bekundungen der Schule verbunden mit dem Ausfall der Fähigkeitsprüfung rechtfertigten in der beträchtlichen Mehrheit der Fälle ein sofortiges eindeutiges Urteil, nämlich überall da, wo diese Kriterien in günstigem oder ungünstigem Sinne zu-

sammenstimmten. Die eigentliche Arbeit des Aufnahmeausschusses setzte bei den zweifelhaften Fällen ein, deren Testausfall sich um die oben erwähnte Grenzpunktzahl herumlagerte.

„Hier mußte sorglich abgewogen werden. Besondere Empfehlung des Lehrers, besonders gute Zeugnisse verhalfen noch manchen Kindern zur Aufnahme, die der Prüfung nach schon etwas unterhalb der Punktzahlgrenze standen. Auch der Beobachtungsbogen gab hier oft recht wertvolle Aufklärungen. Erwähnte er z. B., daß das Kind besonders gründlich und daher etwas langsam arbeite, so konnte eine etwaige Unvollständigkeit einiger Testleistungen, die das Prädikat gedrückt hatten, milder beurteilt werden. Hob der Beobachtungsbogen starkes selbständiges Denken, sinnvolles Fragen, Fähigkeit, das Wesentliche zu erfassen, und ähnliches besonders hervor, so sahen wir darin Äußerungen der spontanen I. und neigten zu positiver Entscheidung, selbst wenn die Äußerungen der gebundenen reaktiven I. bei der Prüfung nicht so günstig ausgefallen waren. In anderen Fällen, die nach dem Testausfall allein zweifelhaft waren, klärte uns der Beobachtungsbogen auf, daß die bisherigen Schulleistungen des Schülers mehr auf rein gedächtnismäßige als auf intellektuelle Funktionen zurückzuführen seien; für ein solches Kind schien der Verbleib in der Normalschule geeigneter. Ferner gab es Schüler, bei denen Schulzeugnisse und Beobachtungsbogen erwiesen, daß ihre Stärke und ihr Hauptinteresse weniger auf sprachlich-geschichtlichem Gebiet als auf dem Gebiet des Rechnens oder der beobachtenden Naturwissenschaften oder der Technik liege; war in solchen Fällen nicht laut Testausfall eine allgemeine I. von ungewöhnlicher Höhe vorhanden, dann schien es nicht richtig, die Kinder in Klassen zu überführen, die hauptsächlich auf sprachlichem Gebiet erhöhte Anforderungen stellen.“[1])

Selbstverständlich waren auch diese Entscheidungen letzten Endes nicht von Willkür frei; aber viel stärker war doch bei dem Aufnahmeausschuß der Eindruck, wie groß der Vorteil war, so viele Urteilsgrundlagen zur Verfügung zu haben; eine ganze Reihe von Schülern wäre zweifellos unrichtig beurteilt und ungerechtfertigterweise aufgenommen bzw. abgelehnt worden, wenn man sich lediglich auf Lehrerurteil und Zeugnis, aber auch wenn man sich lediglich auf den Testausfall verlassen hätte.

Das Verfahren im Jahre 1919 wich in einigen Punkten von dem des Vorjahres ab. Die Zahl der Kinder war noch beträchtlich höher als 1918: 1658 wurden vorgeschlagen; 1260 sollten aufgenommen werden. War es schon im Vorjahre nur mit großer Mühe gelungen, die zur Prüfung von 1355 Kindern nötige gewaltige Organisation zu schaffen und reibungslos arbeiten zu lassen, so schien es ganz unmöglich, nach so kurzer Zwischenzeit einen noch größeren

[1]) *218*, S. 21.

Apparat ins Spiel zu setzen. Es wurde deshalb darauf verzichtet, alle vorgeschlagenen Kinder zu prüfen, und wir glaubten zu diesem Verzicht um so eher berechtigt zu sein, als 1919 die Beobachtungsbogen meist gründlicher ausgefüllt worden waren als im Vorjahr, und daher reichlicheren Stoff zur Beurteilung der Kinder lieferten. Infolgedessen entschloß sich der Aufnahmeausschuß bei 878 Kindern (also etwas mehr als der Hälfte der vorgeschlagenen), ohne Testprüfung lediglich auf Grund der Beobachtungsbogen, Schulzeugnisse und Lehrerempfehlungen die Aufnahme auszusprechen. Die übrigbleibenden 780 wurden geprüft; von diesen konnte noch die Hälfte aufgenommen werden. Die Prüfung war ganz ähnlich organisiert wie im Vorjahre, nur daß sie auf $3\frac{1}{2}$ Stunden eines Vormittags beschränkt war.

Es wurden folgende Tests angewandt:

A. Analogietest[1]). In jeder von 20 Einzelaufgaben sind drei Wörter gegeben, zu denen ein viertes gesucht werden soll, das sich zum dritten verhalten soll wie das zweite zum ersten. Es handelt sich also hier um eine Leistung logischer Verknüpfung. Die Aufgabe wurde den Kindern an folgenden drei Beispielen in durchaus anschaulicher Form nahegebracht:

Lampe — Licht; Ofen — Wärme;
waschen — Gesicht; fegen — Stube;
Kasten — Deckel; Schrank — Tür.

B. Verben-Ergänzungstest. In einer kleinen Erzählung, die nach Inhalt und sprachlicher Form der Auffassung 10jähriger Kinder entspricht, sind 14 Lücken vorhanden. Für jede dieser Lücken sind möglichst viele synonyme Verben zu finden.

C. Bildbeschreibung. Die in einem Bild dargestellte einfache Handlung ist in Form eines Aufsatzes niederzuschreiben.

D. Wortauswahl-Test. Ein Lücken-Ergänzungstest mit Wortliste, die für jede Lücke mehrere Wörter enthält. Das passende Wort ist auszuwählen und in die Lücke einzufügen.

E. Verstandesfragen und Vergleiche. Drei Fragen, in denen Situationen verständig beurteilt werden sollen. Drei Wortpaare, in denen die zugeordneten Glieder verglichen werden sollen.

Die Wertung fand wiederum für jeden Test in Arbeitsgemeinschaften statt unter folgenden Leitern: Test A. Lehrer Rudolf Peter, Test B. Dr. Höper, Test C. Seminarlehrer Penkert, Test D. Dr. H. Werner, Test E. Lehrerin Martha Muchow.

Die Art der Wertung unterschied sich von der vorjährigen dadurch, daß nicht Gesamtprädikate von 1—5 für jedes Test-

[1]) Der Test ist einem ähnlichen, noch nicht veröffentlichten, nachgebildet, der vom Institut f. angewandte Psychologie (Kleinglienicke) für Aufmerksamkeitsprüfungen ausgearbeitet worden war.

ergebnis herbeigeführt, sondern die Rangplätze der Leistungen festgestellt wurden. Für jedes Kind wurden dann die Rangplätze, die es in den fünf Tests erzielt hatte, zu einer Gesamtpunktzahl verbunden, die als endgültiger Wert der Beurteilung zugrunde gelegt wurde.

Die Aufnahme erfolgte auf Grund einer Vergleichung des Prüfungsausfalls mit den sonstigen über das Kind vorliegenden Daten.

Werfen wir nun einen kurzen Blick auf die Ergebnisse und Lehren dieser Veranstaltungen, soweit diese nicht schon in den grundsätzlichen Erörterungen des vorigen Abschnitts vorweggenommen sind.

Zunächst das eine: beide Auslesen sind insofern durchaus gelungen, als sie für die große Masse der in Betracht kommenden Kinder die höchste überhaupt erreichbare Wahrscheinlichkeit richtiger Entscheidung herbeigeführt haben. Jedes andere Verfahren hätte zu viel größeren Unsicherheiten und Fehlgriffen führen müssen. Über die absolute Höhe der prognostischen Sicherheit können freilich erst künftige Jahre Auskunft geben, wenn wir zu überschauen vermögen, wieviel der damals ausgelesenen Kinder in den F-Zügen versagt haben. (Vgl. S. 294: „Bewährung".)

Es dürfen aber andererseits nicht die Schwierigkeiten verschwiegen werden, die einer solch groß angelegten Ausleseveranstaltung entgegentreten. Die Organisation eines so umfangreichen wissenschaftlich begründeten Prüfungsverfahrens verlangt einen Apparat, dessen einzelne Teile mit unbedingter Sicherheit ineinander greifen und dem leitenden Gedanken sich fügen müssen. Es darf als ein besonderes Glück betrachtet werden, daß es 1918 gelang, einen Stab von 60—70 genau vorbereiteten Prüfern und Bearbeitern und die Hilfsmittel des psychologischen Laboratoriums in den Dienst dieser Aufgabe stellen zu können; aber schon 1919 schien eine Einschränkung nötig — wodurch bedauerlicherweise für einen Teil der Kinder auf eine Prüfung verzichtet werden mußte; und die regelmäßige Wiederholung solch umfassender Veranstaltungen scheint in dieser Form unmöglich. Das gilt um so mehr, als in Zukunft aus den 10jährigen Volksschülern nicht nur die Besucher der F-Klassen, sondern auch die künftigen Schüler der höheren Schulen ausgelesen werden müssen, wodurch sich die Zahl der in Betracht kommenden Kinder verdoppelt. Es wird deshalb eine Verfahrungsweise anzustreben sein, die den Auslese-

prozeß in organischerer Form dem natürlichen schultechnischen Betrieb annähert, etwa in der Form, daß die Aufnahmeprüfungen (Leistungsprüfungen) mit Fähigkeitsprüfungen mehr oder minder eng verknüpft werden — wobei die Erfahrungen der bisherigen Auslese natürlich im weitesten Umfange nutzbar gemacht werden können.

Eine zweite Schwierigkeit liegt in dem Verhalten weiter Kreise der Lehrerschaft. Daß und warum diese zum Teil einem psychologisch gerichteten Ausleseverfahren nicht geneigt ist, wurde schon im grundsätzlichen Teil besprochen; dort wurde auch dargelegt, daß die Gründe der Ablehnung nicht stichhaltig sind; und es ist zu hoffen, daß der Zwang der Umstände den Lehrern allmählich zeigen wird, daß ohne die Beteiligung psychologischer Auslesemaßnahmen die gegliederte Einheitsschule überhaupt nicht durchgeführt werden kann. Aber neben der Abneigung kommt auch die fehlende Einstellung der Lehrer auf die neue Aufgabe in Betracht, und dies zeigt sich besonders dort, wo eine unmittelbare Beteiligung des Lehrers beansprucht wird, also beim Beobachtungsbogen.

Die ersten Ausfüllungen im Jahre 1918 sind in nicht seltenen Fällen recht dürftig ausgefallen. Mochte das zum Teil an äußeren Gründen — wie Fehlern in der Fassung der Fragen und zu geringer Frist zwischen Empfang und Ablieferung des Bogens — beruhen, zum größeren Teil lag es doch daran, daß die Herstellung eines solchen psychologischen Rechenschaftsberichts über ihre Schüler für viele Lehrer eine durchaus fremdartige Aufgabe war, mit der sie sich ohne rechte innere Anteilnahme schlecht und recht abfanden. So begegnete man oft in den ausgefüllten Bogen einer überraschenden Einsilbigkeit der Angaben oder einem unbedingten Sichhalten an die vorgeschlagenen Antworten; ja es gab Fälle, in denen die Beobachtungsbogen der sechs von einem Lehrer vorgeschlagenen Kinder so gleichlautend ausgefüllt waren, daß daraufhin überhaupt keine Differenzierung dieser Kinder möglich gewesen wäre. Daneben gab es freilich auch viele, denen man die Gründlichkeit und Liebe bei der Arbeit anmerkte, auch solche, die durch sehr charakteristische Bemerkungen, durch Angaben, die über die Fragen weit hinausgingen, durch beigelegte Proben von Aufsätzen usw. ein außerordentlich anschauliches Seelenbild der Kinder entwarfen, so daß der Aufnahmeausschuß diese Bogen mit großem Nutzen verwerten konnte.

Das Jahr 1919 zeigte schon ein sehr viel erfreulicheres Bild; es war nicht nur die Zahl der besonders gründlich und charak-

teristisch ausgefüllten Bogen beträchtlich gestiegen, sondern auch das Durchschnittsniveau der Bogen war besser geworden. Und so darf die Hoffnung ausgesprochen werden, daß trotz der immer noch lebhaften Widerstände der Beobachtungsbogen allmählich zu einem regelmäßigen und brauchbaren Hilfsmittel der Auslese werden wird, um so mehr, als gerade er ja dem Lehrer die Gewähr gibt, daß sein Urteil die ihm zukommende Bedeutung bei der Entscheidung erhalte.

Zugleich aber lehrt doch diese Ungleichmäßigkeit in der Ausfüllung der Bogen, in wie verschiedenem Maße die Lehrer imstande sind, die Individualitäten ihrer Schüler zu erkennen, und wie notwendig es daher ist, nicht allein dem Lehrer die Entscheidung über die Zukunft der Kinder zu überlassen.

Daß die Testprüfung nicht in jeder Hinsicht einwandfrei war, ist selbstverständlich. Trotz gründlicher Vorproben war gelegentlich bei dem einen oder anderen Test die freigegebene Arbeitszeit zu kurz bemessen worden, so daß ein kleiner Teil der Kinder nicht ganz fertig wurde. Der Kritiktest erwies sich für die geprüfte Altersstufe als zu schwer, so daß die Leistung sich nicht mit der wünschenswerten Deutlichkeit differenzierte. Manche Schwierigkeiten boten auch die Bewertungen der Tests. Es waren die Bewertungsprinzipien so herauszuarbeiten, daß die wesentlichen Seiten der Leistung nach Güte und Umfang darin zum Ausdruck kamen und daß zugleich alle an einem Test Mitarbeitenden durchaus gleichmäßig urteilten; diese Forderung ist um so schwerer durchzuführen, je mehr der Test freie und spontane Betätigung des Kindes zuließ. Darum erwiesen sich Aufsätze über Bilder und Bilderserien — Tests, die auf besonderen Wunsch der Behörde eingefügt wurden und die auch psychologisch besonders interessant sind — als recht spröde gegenüber der ziffernmäßigen Analyse und Einschätzung, wie sie nun einmal durch das Ziel der Veranstaltungen gefordert wurden. Darum werden hier so manche Willkürlichkeiten und Schwankungen mit untergelaufen sein. Aber solche — bei einem ganz neuartigen Verfahren selbstverständlichen — Unvollkommenheiten brauchen uns nicht abzuschrecken; sie sind vielmehr die wichtigsten Lehrmeister für künftige Ausgestaltungen und Verbesserungen der Methoden.

Psychologische Ergebnisse. Abgesehen von dem rein praktischen Zweck der Auslese stellen die Beobachtungsbogen und Test-

niederschriften von mehreren Tausend begabter 10jähriger Kinder ein in ihrer Weise einzigartiges Material für psychologische Untersuchungen der mannigfaltigsten Art dar.

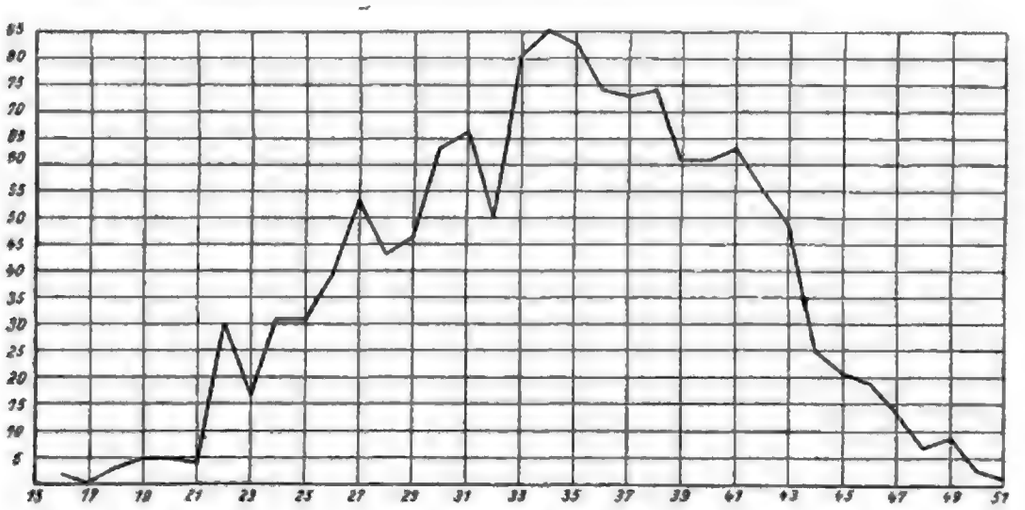
Leider war es bisher nur in ganz beschränktem Maße möglich, solche Bearbeitungen in die Wege zu leiten; und das, was bis jetzt aus diesem Material entnommen werden kann, bezieht sich mehr auf Problemstellungen und Gesichtspunkte als auf positive Ergebnisse. Daher muß an dieser Stelle folgendes genügen.

Die ausgefüllten Beobachtungsbogen ermöglichen eine doppelte Verarbeitung: im Sinne einer differentiellen Psychologie begabter Schüler und einer der beschreibenden Lehrer. Was das erste anlangt, so erlaubt eine massenstatistische Bearbeitung eine Beantwortung der Frage, welche Eigenschaften bei allen (oder fast allen) vorgeschlagenen Schülern stark ausgeprägt sind, und in welchen Eigenschaften größere Variationen zwischen ihnen bestehen. Die ersten Eigenschaften werden diejenigen sein, auf welche die Lehrer vor allem ihr Urteil über höhere Begabung gründen. Auch Korrelationen zwischen bestimmten Eigenschaften — ferner Unterschiede der Geschlechter — werden sich aus den Psychogrammen ableiten lassen. Bei den Ausfüllern hinwiederum werden sich Grade und Typen der psychologischen Beobachtungs- und Urteilsfähigkeit herausstellen, wodurch das bisher so stiefmütterlich behandelte und doch so wichtige Gebiet der Psychologie des Lehrers ganz neue und zum Teil recht überraschende Aufhellungen erfährt.

Die Testprüfung legt zunächst die Untersuchung der einzelnen intellektuellen Funktionen nahe, die sich in den Testleistungen betätigen. Eine solche Analyse ist bereits mit manchem Erfolg von Herrn Meins (*218*) für den Kritiktest durchgeführt worden; andere Analysen stehen leider noch aus. Ebenso harrt noch die weitere Frage der Beantwortung, in welcher Korrelation die einzelnen Testleistungen untereinander stehen; der große Umfang des Materials würde hier Korrelationsbestimmungen von sonst kaum gekannter Zuverlässigkeit erlauben und daher auf die größere oder geringere Engigkeit der Beziehung zwischen verschiedenen intellektuellen Leistungen Licht werfen.

Daß im ganzen die geprüften Funktionen in einer nicht zu geringen Korrelation zueinander stehen müssen, läßt sich allerdings indirekt aus der Häufigkeitsstatistik der Prädikate ableiten. Diese ist in der beiliegenden Kurve dargestellt. Die Abszissen entsprechen den Summen der Punkte, die die einzelnen Kinder

erreichten; die Ordinaten geben die Anzahl der Kinder, die die durch die Abszisse ausgedrückte Punktzahl erhalten haben. Die Kurve ist, wie man sieht, im wesentlichen eingipflig symmetrisch; sie entspricht ungefähr der normalen Gaußschen Verteilungskurve (s. S. 158). Die Symmetrie der Kurve ergibt sich mit Selbstverständlichkeit aus den Voraussetzungen; denn für jeden Test sollten die Prädikate 1—5 in ungefähr symmetrischer Verteilung verwandt werden; daher müssen sich auch die Gesamtprädikate, die durch Summation aus jenen Einzelprädikaten entstanden, symmetrisch verteilen. Die Gaußsche Verteilung dagegen zeigt, daß die Komponenten, aus denen die Gesamtprädikate gebildet wurden, in merklicher Korrelation zueinander stehen müssen. Denn wenn die ganz niedrigen und die ganz hohen Punktziffern in der zu erwartenden



Häufigkeit vorkommen, so bedeutet dies, daß die betreffenden Individuen in allen oder fast allen Tests gleichmäßig gute bzw. schlechte Prädikate erhalten haben müssen. Ständen die Tests in gar keiner Korrelation zueinander, d. h. würden sich gute und schlechte Einzelprädikate bei allen Individuen beliebig gestreut finden, dann würden sich alle Gesamtprädikate außerordentlich ähneln; die Häufigkeitskurve würde nur aus ganz wenigen mittleren Werten bestehen und des allmählichen Abfalls nach oben und unten entbehren.

So scheint denn auch die Massenstatistik unsere Auffassung der „Intelligenz" als einer in sich zusammenhängenden, über verschiedene Leistungsformen sich erstreckenden Leistungsfähigkeit zu bestätigen.

Ein Vergleich der Ergebnisse für beide Geschlechter ergibt folgendes. 1918 bestanden keine exakten Vergleichsgrundlagen,

denn es waren von den Lehrern entsprechend der größeren Zahl einzurichtender Knabenklassen 882 Nennungen von Knaben erfolgt gegenüber nur 473 Mädchen. Es war also bei den Mädchen die Auslese eine viel schärfere gewesen, und es ist deshalb selbstverständlich, daß der Gesamtdurchschnitt der Testleistungen bei den Mädchen ein wenig über der der Knaben steht (Mädchen 33,4, Knaben 35 Punkte; die höhere Punktzahl bedeutet bekanntlich die geringere Leistung). Will man eine ungefähre Gleichartigkeit der Bedingungen herbeiführen, so muß man die 473 Mädchen mit jenen Knaben vergleichen, die von den Lehrern vorgeschlagen wären, wenn sie nur 400—500 hätten vorschlagen dürfen. Das wären offenbar im großen und ganzen diejenigen Knaben gewesen, deren Namen von den Lehrern — als Zeichen besonderer Empfehlung — bei der Nennung unterstrichen worden waren; es sind 421. Für diese beträgt die Durchschnittspunktzahl 33,2; sie deckt sich fast mit der oben genannten Mädchenpunktzahl 33,4.

Im Jahre 1919 lagen die Verhältnisse für einen Geschlechtervergleich besonders günstig, da gleich viele Knaben und Mädchen von den Lehrern vorgeschlagen wurden, und von diesen vorgeschlagenen wiederum eine gleiche Anzahl (nämlich je 390) zur Prüfung kam. Bisher war es leider nur möglich, für einen Test die Vergleichsrechnung durchzuführen[1]). Es war der Analogietest (s. S. 288), bei dem 20 Aufgaben vorgelegt wurden. Es ergab sich nun als durchschnittliche Zahl richtiger Lösungen bei den Knaben 14,6, bei den Mädchen 14,3, also wieder fast völlige Deckung[2]).

Die Bewährung der Hamburger Schülerauslese. Nach Fertigstellung des obigen Berichtes erscheint eine Veröffentlichung, die im Auftrage der Hamburger Oberschulbehörde nach amtlichem Material von Dr. Höper zusammengestellt ist (*191*a). Sie enthält zahlreiche Tabellen über die an den ausgelesenen Schülern bisher gemachten Schulerfahrungen und erlaubt wichtige Rückschlüsse auf die Brauchbarkeit der Hamburger Auslesemethode. Es ergibt sich, daß sich das Verfahren in seiner gemeinsamen Benutzung von Lehrerurteil und Fähigkeitsprüfung recht gut bewährt hat, weit besser als das Berliner Verfahren.

[1]) Ausgeführt von Herrn R. Peter.

[2]) Die nähere Analyse zeigt allerdings, daß der kleine Vorsprung der Knaben vor allem durch ihre bessere Leistung bei den beiden schwersten Aufgaben bedingt ist; sie lauten: Wasser, Schiff; Land, —; Haus, Hütte; Fluß, —. Dort stammen 60%, hier 53% aller richtigen Lösungen von den Knaben.

Besonders ergiebig sind die Befunde, die über die Ostern 1918 ausgelesenen Schüler vorliegen, weil sich diese Ergebnisse bereits auf 1½jährige Schulerfahrungen stützen können. Eine Umfrage der Behörde bei den Lehrern der F-Klassen verlangte Zeugnisprädikate über Begabung, Fleiß, Leistungen in Sprachen und exakten Fächern, ferner Angaben über die Zahl der Schüler, die bereits wieder die F-Klasse verlassen mußten und über die Gründe dieser „Abschulung".

Die Abschulungen sind für uns am wichtigsten, weil sie den negativen Maßstab für die Bewährung der Auslese bilden. Für diejenigen Abgänge, die durch Tod, Krankheit, Fortzug der Eltern und ähnliches bedingt waren, ist die Auslese natürlich nicht verantwortlich zu machen (es sei denn, daß man für künftige Fälle die Mitwirkung des Schularztes verlangen muß, der körperlich schwächliche Kinder besonders zu begutachten hätte). Für die psychologische Auslese kommen nur die aus psychologischen Gründen, d. h. wegen Unfleißes oder Begabungsmangels Abgeschulten in Betracht, das sind 43 von 942 oder 4½%. Diese Ziffer ist außerordentlich gering; denn daß von je einer Klasse von 45 Schülern im Durchschnitt 2 infolge unzutreffender Prognose die Anforderungen nicht erfüllen, ist bei der Unberechenbarkeit geistiger Entwicklung eine recht kleine Unstimmigkeit.

Welcher Teil der Auslese ist nun aber irrig gewesen? Unter den 43 Unfleißigen oder Unbegabten waren nicht weniger als 25, also mehr als die Hälfte, die seinerzeit von den abgehenden Lehrern durch Unterstreichen der Namen besonders empfohlen worden waren — schlagender konnte unsere These, daß die Lehrerempfehlung allein zur richtigen Auslese nicht ausreiche, gar nicht bewiesen werden. Die Begabungsprüfungen zeigen eine ganz andere Bewährung. Unter den Abgeschulten sind 25 Kinder, die wegen Begabungsmangels für die F-Klassen ungeeignet waren; das bedeutet aber nicht, daß bei allen diesen die Testprüfung ein falsches Bild gegeben hatte, denn 9 unter diesen 25 hatten in der Prüfung 38 und mehr Punkte, d. h. so schlechte Testleistungen, daß sie auf Grund dieser allein nicht aufgenommen worden wären; nur die Rücksicht auf Beobachtungsbogen und besondere Empfehlungen des Lehrers hatte ihnen noch die Zulassung erwirkt. Es bleiben im ganzen 16 Kinder mit genügenden Testergebnissen[1]),

[1]) Unter den Kindern mit guten Testergebnissen brauchte kein einziges abgeschult zu werden.

die sich in der Schule als nicht ausreichend begabt erwiesen. 16 von 942 aufgenommenen Kindern = $1^2/_3\%$ — diese geringe Versagerziffer nach $1^1/_2$jähriger Erprobung zeigt also, daß die Testauslese weit zuverlässiger gewesen war als nicht nur die Gegner, sondern auch die Freunde des psychologischen Verfahrens vermuten konnten [1]).

4. Aufnahmeprüfungen für die höheren Schulen („Mittelschulen") Deutsch-Österreichs (10jährige Kinder).

Der kühne — meines Erachtens verfrühte — Versuch, ein psychologisch-pädagogisches Ausleseverfahren für ein ganzes Land einheitlich zu regeln, ist 1919 in Deutsch-Österreich gemacht worden, und zwar für den Übergang der Volksschüler und -schülerinnen in die unterste Klasse der höheren Schulen (*245, 181, 237*). Das Unterrichtsministerium ordnete unter ausdrücklichem Hinweis auf die Ergebnisse der modernen Psychologie folgendes an:

„Die Auswahl der sich zur Aufnahme in die unterste Mittelschulklasse meldenden Schüler(innen) hat zu erfolgen

a) auf Grund von Schülerbeschreibungen, die über die beim Schüler während des Volksschulunterrichtes gemachten Beobachtungen Aufschluß geben;

b) auf Grund einer Aufnahmeprüfung, durch die nicht bloß das Ausmaß der erworbenen Kenntnisse, sondern hauptsächlich die Begabung des Schülers festgestellt werden soll."

Für die Schülerbeschreibung wird ein Schema vorgelegt, das, wie es scheint, dem Hamburger Beobachtungsbogen nachgebildet ist; doch haben Kürzungen und Änderungen stattgefunden. Über die Auf-

---

[1]) Ähnlich scheinen die Verhältnisse für die Auslese 1919 zu liegen, doch kann man hier noch nicht — bei der erst halbjährigen Erprobungszeit in den F-Klassen — von sicheren Ergebnissen sprechen. Aber eines ist doch bemerkenswert. 1919 war bekanntlich nur die Hälfte der Kinder geprüft worden: die andere Hälfte war ungeprüft aufgenommen worden, weil ihre Eignung nach Lehrerurteil und Schulzeugnis fraglos schien. Nun mußten schon nach einem halben Jahr 6 Mädchen wegen Unfleißes oder Unbegabtheit abgeschult werden; von diesen waren 4 ohne Prüfung aufgenommen worden! (Während die 2 anderen bei der Prüfung nur soeben durchgeschlüpft waren). Bei weiteren 62 Kindern hegen die jetzigen Lehrer starke Zweifel, ob sie nach Fleiß oder Begabung den Anforderungen der F-Klassen entsprechen werden; unter diesen sind wieder 24 ungeprüfte, d. h. nach dem damaligen Lehrerurteil zweifellos geeignete Kinder!

nahmeprüfung, die von den Lehrern der aufnehmenden Schule abzuhalten ist, heißt es:

„Da sich die Aufnahmeprüfung nicht bloß auf die Feststellung eines bestimmten Ausmaßes von durch den Volksschulunterricht erworbenen Kenntnissen beschränken, sondern die für den Eintritt in die Mittelschulstudien ausreichende Begabung des Schülers feststellen soll, sind bei der Aufnahmeprüfung einerseits die Aufgaben und Fragen nach psychologischen Gesichtspunkten auszuwählen und andererseits die Leistungen der Prüflinge entsprechend dem gegenwärtigen Stande der Begabungsdiagnostik zu beurteilen.

Es kommen daher in erster Linie Aufgaben und Fragen in Betracht, die in ihren Lösungen Art und Grad der Begabung (Auffassung, Gedächtnis, Arbeitsart, sprachliche Ausdrucksfähigkeit usw.) deutlich erkennen lassen."

Die erste derartige Aufnahmeprüfung hat im Herbst 1919 stattgefunden, über ihren Verlauf und etwaige Ergebnisse ist noch nichts Näheres bekannt geworden.

Mir scheint, daß man hier die Einführung der Psychologie in die Pädagogik am unrichtigen Ende angefaßt hat. Die verantwortliche Abfassung von Schülercharakteristiken und die Veranstaltung psychologisch orientierter Prüfungen setzt eine psychologisch eingestellte Lehrerschaft und die Mitarbeit eines Stabes geschulter Jugendpsychologen voraus; beides wäre in Deutsch-Österreich erst in geduldiger Arbeit zu schaffen. Zudem ist der methodische Stand der Begabungsauslese zur Zeit ein solcher, daß man wohl an einzelnen Orten Erprobungen vornehmen darf, um das Verfahren allmählich leistungsfähig zu machen, nicht aber einem ganzen Lande schon ein festes Schema auf dem Verordnungswege auferlegen kann. Man darf sich also nicht wundern, wenn der Versuch in dieser Form einen Schlag ins Wasser bedeuten sollte; man braucht sich aber dadurch auch nicht entmutigen zu lassen.

5. Prüfung 11jähriger Knaben (Göttingen).

Der Göttinger Lehrer Kaufel (*194*) erhielt vier von den Lehrern ausgelesene begabte Knaben zur Ausbildung in einem Förderkursus, der sie auf die Mittelschule vorbereiten sollte. Da K. sich nicht allein auf die Unterrichtsleistungen und Schulbeobachtungen verlassen wollte, stellte er eine Fähigkeitsprüfung an, die sich im wesentlichen an die Berliner Methode anlehnte

(s. den folg. Abschnitt); an einigen Stellen waren mit Rücksicht auf das jüngere Alter einfachere Tests gewählt worden. Es ergab sich auf Grund der Fähigkeitsprüfung dieselbe Rangordnung wie nach den Schulergebnissen, so daß nun mit gutem Gewissen der vierte ausgeschieden, der dritte einer etwas niederen Klasse als der erste und zweite zugewiesen werden konnte.

6. Auslese 13jähriger Schüler und Schülerinnen (Moede und Piorkowski, Berlin).

Die Berliner Begabtenauslese[1]) ging zeitlich allen anderen voraus, und da sie durch eine populär gehaltene Darstellung sofort in weitesten Kreisen bekannt wurde, hat sie lange Zeit so gut wie ausschließlich die Erörterung über das ganze Problem beherrscht. Dies ist insofern bedauerlich, als in der breiten Öffentlichkeit die Meinung entstehen konnte, daß die dort vertretenen psychologischen Anschauungen und angewandten Methoden den Standpunkt der Psychologie zum Problem der Begabtenauslese darstellen. In Wirklichkeit aber sind die kritischen Stimmen nicht etwa nur auf der Seite derer laut geworden, die die Beteiligung der Psychologie bei der Auslese überhaupt ablehnen, sondern auch innerhalb der Psychologie selbst, die gegen diese Form psychologischer Mitarbeit ernste Bedenken erhob[2]).

---

[1]) Wir nennen hier nur die wichtigste Literatur. Das erste Buch von Moede-Piorkowski-Wolff (*206*) liegt bereits in dritter Auflage vor; der neuesten Auflage ist eine — auch gesondert erschienene (*207*) — Zurückweisung der bisherigen Kritiken beigegeben. Der Nachweis, daß die Bewährung der Ausgelesenen nicht ganz den Erwartungen entspricht, ist von Schönebeck (*223—225*) geführt worden, dem dann Moede und Piorkowski ihrerseits erwidern (*208*). Angekündigt ist eine Darstellung der praktischen Schulerfahrungen an den ausgelesenen Schülern (*209*).

[2]) Eine eingehende wissenschaftliche Begutachtung des Berliner Verfahrens wird leider noch dadurch erschwert, daß die beiden daran beteiligten Psychologen Moede und Piorkowski uns noch immer nicht die dringend erforderliche dokumentarische Darstellung ihrer Untersuchungsmethodik und des Bewertungs- und Verrechnungsverfahrens gegeben haben. Wir anderen Psychologen, die wir uns an der Begabungsauslese beteiligen, sind der Überzeugung, daß sich erst ganz allmählich eine vollbefriedigende Methodik aus den immer wieder zu erneuernden Erfahrungen entwickeln werde, und halten es daher für erforderlich, unsere Methoden mit der Objektivität und Exaktheit zu schildern, die auch anderen erlaubt, sich über Vorzüge und Schwächen eine eigene Überzeugung zu bilden. M. und P. scheinen dagegen der Meinung zu sein, daß ihre Testzusammen-

In Berlin wurden 1917 höhere Schulen mit verkürztem Lehrgang eingerichtet, die für hochbefähigte Gemeindeschüler des siebenten Schuljahres bestimmt waren. Das verkürzte Gymnasium und Realgymnasium setzt mit einer für beide Richtungen gemeinsamen Tertia ein, von Sekunda an findet die Zweiteilung statt; die Schule ist bestimmt für solche hochbefähigte Schüler, die sich ihrer ganzen Anlage nach zum Studium eignen. Weiter wurde eine verkürzte Realschule gegründet, die in dreijährigem Kurs die Ziele der Realschule erreichen und eventuell zur Oberrealschule überleiten soll. In diese Schule werden hochbefähigte, zum Studium der technischen Fächer geeignete Gemeindeschulknaben aufgenommen. Endlich wurde eine Mädchenmittelschule neu geschaffen, zu deren „Förderklassen" Volksschülerinnen im Alter von 12—13 Jahren übertreten sollten.

Die für diese Schulen vorzunehmende Auslese sollte sehr scharf sein, so daß wirklich nur die allerhöchsten Begabungen hineinkommen könnten. So wurde für das Begabtengymnasium ein jährlicher Zugang von höchstens 90 Knaben vorgesehen, das sind 1½% der 6000 männlichen Volksschüler des siebenten Schuljahres. Das Ausleseverfahren war nun für das Begabtengymnasium in Kürze das folgende (für die anderen Schulen entsprechend):

Die Rektoren der Berliner Gemeindeschulen wurden aufgefordert, diejenigen Schüler der ersten Klasse vorzuschlagen, die ihnen nach der Höhe ihrer Begabung für die Aufnahme in eine Begabtenschule geeignet erschienen. Nähere Angaben, nach welchen die Rektoren die Begabungen zu beurteilen hatten, wurden nicht gemacht; Beobachtungsbogen wurden nicht ausgefüllt[1]). Diese vorgeschlagenen Knaben — deren Gesamtzahl nicht genannt wird —

---

stellung bereits ein fertiges und treffliches Verfahren darstelle, das von anderen auf Grund der ungefähren Andeutungen ihres Buches übernommen werden kann. Es steckt gewiß sehr viel Gutes in dem Berliner Verfahren; und der kühne Wurf, mit dem die beiden Psychologen für eine durchaus neuartige Aufgabe in kürzester Frist eine Methode schufen, verdient volle Anerkennung; aber die Selbstsicherheit, mit der sie an die Güte ihrer Methode glauben, ist verfrüht; und wenn sich jetzt herausstellt, daß der Methode beträchtliche Irrtümer unterlaufen sind, so trägt die erste allzu siegesgewisse Darstellung die Hauptschuld an der Enttäuschung, die nun leicht entstehen kann und die nicht nur die Berliner Auslese, sondern auch — ungerechtfertigterweise — die Psychologie überhaupt in Mißkredit zu bringen vermag.

[1]) Neuerdings ist der Beobachtungsbogen auch in Berlin eingeführt, und zwar in der Rebhuhnschen Fassung (*221*).

wurden nun einer Prüfung unterzogen, die allein von den beiden Psychologen, ohne bestimmende Mitwirkung von Pädagogen, vorbereitet, angestellt und ausgewertet wurde. Die Prüfung war als „Funktionsprüfung" gedacht, d. h. es sollte systematisch der ganze Umkreis der in Betracht kommenden psychischen Funktionen in ihren Hauptleistungen der Analyse unterworfen und durch Kombination der Einzelleistungen für jeden Schüler das intellektuelle Gesamtverhalten einwandfrei rangiert werden (*206*, S. 109). Die Ausarbeitung einer solchen Prüfung kann, wie die Prüfer meinten, „nicht schwer fallen" (S. 110) — in Wirklichkeit ist es eine Aufgabe von ungeheurer Schwierigkeit, die, wenn man sie wirklich „systematisch" und „einwandfrei" vorbereiten wollte, Jahre erfordern würde. Was in den wenigen Wochen der verfügbaren Vorbereitungszeit möglich war, beschränkte sich auf die Zusammenstellung und Abänderung einiger von früher her bekannter Tests und der Hinzufügung einiger ad hoc konstruierter neuer, die nun aber ohne genaue Vorprüfung und Eichung angewandt werden mußten. Es soll nicht bestritten werden, daß die beiden Forscher hierbei teilweise eine glückliche Hand zeigten und unseren Methodenschatz um aussichtsvolle Verfahrungsweisen vermehrten; dennoch bleibt bestehen, daß die Testzusammenstellung eine geschickte Improvisation war und auch infolge der drängenden Umstände nichts anderes sein konnte; es hätte nichts geschadet, wenn dies offen ausgesprochen worden wäre.

Über die Gesamtorganisation der Prüfung erfahren wir aus dem Bericht nur, daß in Gruppen geprüft wurde, und daß die Prüfung etwa 10 Stunden dauerte, die auf mehrere Tage verteilt waren. Über die Größe der Gruppen, die Gesamtzahl der zu Prüfenden und die der auszulesenden Kinder, die Einteilung der Prüfung usw. fehlt jegliche Angabe. Auch darüber wird nichts gesagt, welche Vorsichtsmaßregeln getroffen waren, um einen einwandfreien Verlauf der Prüfung zu sichern, um Störungen durch unbeteiligte Zuschauer und das gegenseitige Absehen und Vorsagen der Prüflinge zu vermeiden. Eigener Augenschein (ich wohnte einmal einer Prüfung bei) und Angaben der Literatur (z. B. von Schönebeck) beweisen, daß hier nicht alles so war, wie es hätte sein müssen.

M. und P. stellen das folgende Schema ihrer Prüfung auf (in Klammern ist von uns stets die angewandte Methode angedeutet):

I. Aufmerksamkeits- und Konzentrationsfähigkeit bei unmittelbarem und reproduktivem Material.
   a) Dauerspannung (Durchstreichen von a, e und n in vorgelegten Texten), (Hinschreiben von Worten, die sich auf Gegenstände des Zimmers beziehen, zweisilbige Dingwörter sind und kein a, e, n enthalten).
   b) Ablenkbarkeit und Mehrfachhandlungen (während im Kopf multipliziert wird — nur die Endergebnisse sind hinzuzuschreiben —, muß zugleich auf eine vorgelesene Geschichte geachtet werden, die nachher aus dem Gedächtnis niederzuschreiben ist).
   c) Ermüdbarkeit (Wiederholung des Durchstreichversuches nach mehrstündiger Prüfungsarbeit).

II. Gedächtnis:
   A. Zuführung neuen Gedächtnisstoffes.
      a) Gedächtnis für sinnlosen Stoff (Trefferversuche mit Paaren sinnloser Silben und mit Wort-Ziffer-Paaren).
      b) Gedächtnis für sinnvolle Stoffe (Trefferversuche mit dreigliedrigen Kausalzusammenhängen, z. B.: Arbeit — Verdienst — Wohlstand usw.).
   B. Bestand der vorhandenen Dispositionen, ihre Bereitschaft und Abwicklung („Themenassoziation": auf ein gegebenes Wort hin mußten die Kinder alles niederschreiben, was ihnen einfiel).

III. Kombination.
   A. Anschauliche Kombination (die Teile eines zerschnittenen Rechtecks wurden gezeigt; die Figur mußte in richtiger Zusammensetzung aufgezeichnet werden).
   B. Intellektuelle Kombination.
      a) Gebunden (Ergänzung eines Lückentextes).
      b) Frei (aus drei gegebenen Begriffen möglichst viele verschiedene sinnvolle Zusammenhänge herstellen. Sogenannter Variationsversuch).

IV. Begriffsbereich.
   A. Bestand an vorhandenen Begriffen und seine Flüssigkeit (Definitionen).
   B. Stiftung neuer begrifflicher Beziehungen.
      a) Heraussuchen des Wesentlichen unter gegebenen Elementen (knappe Angaben des Hauptinhalts einer ausgeschmückten Geschichte).
      b) Finden des Gemeinsamen zwischen gegebenen Gliedern einer Reihe (verschiedene geometrische Figuren).
      c) Erfassen funktionaler Beziehungen zwischen mehreren Merkmalsreihen (ebenfalls an geometrischen Figuren).

V. Urteilsfähigkeit.
   A. Allgemeine Beurteilung auf Grund
      a) sachlicher Wertung der Umstände (Beurteilung eines Schlachtberichts);
      b) seelischer Einfühlung in Wirklichkeits- und Bildversuch sowie bei sprachlicher Darbietung (Deutung eines dramatischen Bildes. Weiterführung einer ohne Abschluß abgebrochenen Erzählung);
      c) sachlich-psychologischer Wertung des Tatbestandes (Beurteilung eines moralstatistischen Problems).

B. Beurteilung von Sonderfällen.
   a) Erfassen des Wahrscheinlichsten bei gegebenen Umständen in dargebotenen Beispielen.
   b) Finden des Zweckmäßigsten in einer gegebenen Situation.

VI. Anschauungs- und Beobachtungsfähigkeit.
   A. Anschauungsfähigkeit (aus gefaltetem Papier wird ein Dreieck ausgeschnitten; welche Figur entsteht?).
   B. Beobachtungs-Schärfe und -Ergiebigkeit.
      a) Aussage über Dinge und Merkmale im Bildversuch (Bericht und Verhör mit Suggestivfragen).
      b) Erfassen von Relationen in der Wahrnehmung (an einfachen technischen Modellen sollen die daran möglichen Bewegungsvorgänge aufgefaßt werden).

Auf eine kritische Besprechung der Einzelmethoden muß an dieser Stelle verzichtet werden, um so mehr, als die Angaben über die Methodik, die den Kindern gegebenen Anleitungen und die Art der Bewertung zu einer Nachprüfung unzureichend sind. Statt eines systematischen Berichts werden uns nur beliebig herausgegriffene Beispiele von Aufgaben und Lösungen gegeben und einige Betrachtungen über Geschlechtsunterschiede und andere Ergebnisse hinzugefügt, für die uns wiederum die Unterlagen vorenthalten werden.

Für jeden einzelnen Test wurden die Leistungen quantitativ gewertet und hiernach in eine Rangordnung gebracht. Sodann wurden innerhalb jeder mit römischer Ziffer bezeichneten Testgruppe die Rangordnungen zu einer gemeinsamen Rangordnung verbunden. Endlich wurden diese sechs Rangordnungen wiederum zu einer Gesamtrangordnung vereinigt. Der Rangplatz, den jedes Kind in dieser Gesamtrangordnng erhielt, war maßgebend für die Entscheidung über Aufnahme oder Ablehnung.

Nun ist die Herstellung eines solchen Gesamtprädikates eine unumgängliche Maßnahme, wenn man überhaupt aus der Gesamtheit der Prüfung eine einheitliche Resultante ziehen will. Aber der Resultantenwert ist um so weniger ausreichend, je heterogener die in ihm versteckten Einzelleistungen sind. Darum erlaubt eine „allseitige geistige Funktionsprüfung", wie sie von M. und P. angestrebt wurde, viel weniger die Zusammenfassung in eine Ziffer, als die „Intelligenzprüfung", die bei der Hamburger Auslese angewandt wurde. An dem Gesamtprädikat der Berliner Prüfungen ist eine so unnatürliche und lebensfremde Leistung, wie die Aufmerksamkeitszerteilung auf zwei disparate gleichzeitige Aufgaben,

und eine so gänzlich unintellektuelle Leistung, wie das Behalten sinnloser Silben, mit gleichem Gewicht vertreten, wie die so wichtigen Leistungen des Kombinierens und Urteilens, oder wie die durchaus auf Spezialbegabung beruhenden Fertigkeiten in visueller Anschauung und technisch-konstruktiver Auffassung. Es scheint nicht einmal dafür gesorgt worden zu sein, daß bei der Aufnahme in die Realschule die Leistungen in den technisch-anschaulichen Tests, bei der Aufnahme in das Gymnasium die Leistungen in den sprachlich-logischen Tests besonders ins Gewicht fielen — vielmehr wurden alle diese qualitativen Verschiedenheiten verwischt durch die rein mechanische Addition sämtlicher Einzelrangplätze.

Die so gewonnenen Ziffern hielten M. und P. für exakt genug, um durch bloße Abzählung den oberen Teil der entstehenden Rangordnung für die Aufnahme zu bestimmen. Von diesem Grundfehler der Berliner Methode ist ja schon mehrfach die Rede gewesen (S. 272/3). Die Psychologen hielten sich hier einfach an die Forderung der Behörde, daß so und so viele Kinder benannt werden sollten; sie zählten diese Kinder ab, ganz gleich, ob ein Teil von ihnen der absoluten Leistungsfähigkeit nach überhaupt für eine höhere Schulbahn ungeeignet schien, ganz gleich, ob vielleicht zwischen dem xten (noch angenommenen) und x + ersten (abgewiesenen) Kind nur ein ganz geringfügiger Unterschied des Gesamtprädikates bestand.

Es ist nicht verwunderlich, daß unter solchen Umständen die Zahl der Fehlurteile weit größer war, als es allein aus der immanenten Unvollkommenheit psychologischer Methodik und aus den unvorhersehbaren Schwankungen der geistigen Entwicklung abzuleiten wäre.

Wie Oberlehrer Schönebeck überzeugend nachweist (*223, 224*), besteht zwischen den nun vorliegenden praktischen Erfahrungen von zwei Schuljahren der Begabtenschule und den Ergebnissen der Fähigkeitsprüfung durchaus nicht die Übereinstimmung, die M. und P. seinerzeit mit Bestimmtheit voraussagten und auch jetzt noch zu sehen glauben.

Von den 34 Schülern einer auf Grund der M. und P.schen Abzählung zusammengestellten Begabtenklasse hatten nach einem Jahr 10 Schüler versagt, so daß sie aus der Klasse ausgeschieden werden mußten. Und für diese Ausscheidungen trifft nicht etwa die Voraussage zu: „Werden einige der ausgelesenen Schüler versagen, so werden dies diejenigen sein, die die untersten Plätze der Fähig-

keitsrangreihe innehatten." Vielmehr ergibt die vergleichende Betrachtung von Schönebeck folgendes: Teilte man die Schüler einerseits nach den Testprüfungen, andererseits nach den Schulleistungen in sehr gute, gute und schwache, so ergab sich, daß unter den nach M. und P. sehr guten 7 Schülern einer so schwach war, daß er die Begabtenklasse aufgeben mußte; von den 18, die bei der experimentellen Leistung gut abschnitten, waren 8, also fast die Hälfte, in der Schule schwach, von diesen 3 geradezu untauglich; umgekehrt waren unter den 9 schwachen Testarbeitern 2, die sich als sehr gute Schüler der Begabtenklasse erwiesen.

Blicken wir zurück und fassen wir zusammen, so läßt sich über die Berliner Begabungsauslese dieses sagen: Ihre eigentliche Bedeutung liegt darin, daß sie Pionierarbeit leistete, indem sie unbeschwert von Bedenklichkeiten der Psychologie den Zugang zu einem bisher unbetretenen Gebiet verschaffte. So gab sie unserer weiteren Arbeit vielerlei Anregungen: wie man es machen muß — und wie man es nicht machen soll! Groß war auch die Propagandawirkung, die von ihr ausging; aber diese war doch eine sehr zwiespältige; denn die Art der Darstellung, zunächst verblüffend und fortreißend, hat doch mit fortschreitender Zeit mehr und mehr jene Schwächen hervortreten lassen, die einem berechtigten Widerstand als Waffe dienen konnten, und so wurde gerade durch das Buch im ganzen unserer praktisch-psychologischen Arbeit so manche Schwierigkeit bereitet, die sachlich vermeidbar gewesen wäre.

7. Die Leipziger Anweisungen (*197*).

Im Anschluß an die Prüfungen von Moede und Piorkowski hat ein Ausschuß des Leipziger Lehrervereins Anweisungen für die psychologische Auswahl jugendlicher Begabter ausgearbeitet. Diese Unternehmung füllt eine empfindliche Lücke aus, die die bisherigen Darstellungen von M. und P. gelassen hatten: es wird nämlich für jeden vorgeschlagenen Test eine genaue Gebrauchsanweisung und eine Methode der Bewertung gegeben. Die Einteilung der Tests deckt sich ganz, die Auswahl der Tests zum größeren Teil mit der von M. und P.; doch begegnen uns einige wertvolle neue Tests, so die „wirren Gedanken" (Ordnen durcheinander gewürfelter Sätze zu einer zusammenhängenden Geschichte), der Telegrammtest zum Herausfinden des Wesentlichen (der Inhalt eines Briefes soll zu einem Telegramm verdichtet wer-

den), der Rechenkritiktest (aus verschiedenen Rechenaufgaben sind unlösbare herauszufinden). Alle Tests sind zwar zum Zweck der Eichung durchgeprobt worden, ihre Anwendung zu Zwecken der Auslese ist aber noch nicht erfolgt.

8. Prüfung 13- und 14jähriger Volksschulkinder in Hannover[1]).

In engem Rahmen bewegten sich die bisherigen Ausleseveranstaltungen in Hannover, da zu jedem Ostertermin nur je eine Förderklasse für begabte Knaben und Mädchen zu schaffen war. Die Mädchen sollen in drei Jahren das Abschlußzeugnis des Lyzeums, die Knaben in gleicher Zeit die Reife für die Unterprima erhalten. Lediglich vorbereitenden Charakter hatte die erste 1918 angestellte Prüfung, die von Dr. Brahn, Leipzig, geleitet wurde. Die eigentlichen Auslesen fanden zu Ostern 1919 und 1920 statt. Da die Zahl der vorgeschlagenen Schüler und Schülerinnen die der aufzunehmenden nicht allzu sehr übertraf, so handelte es sich nur darum, die offenbar Ungeeigneten abzulehnen. 1919 wurden von 16 Knaben 1, von 23 Mädchen 2 ausgeschieden; 1920 von 12 Knaben 0, von 30 Mädchen 8. Beobachtungsbogen wurden nicht verwandt. Die Testprüfungen bezogen sich auf: Aufmerksamkeit, Fähigkeit des Merkens, Beobachtens, Kombinierens, Urteilens, Definierens. Es kamen die gleichen oder analoge Tests zur Anwendung, wie sie von Moede-Piorkowski und in der Leipziger Anweisung beschrieben sind.

9. Auslese von Handelsschülern (Beger, Leipzig [*182*]).

Ebenfalls an das Berliner Verfahren lehnt sich eine kleine Ausleseveranstaltung an, die an der öffentlichen Handelslehranstalt in Leipzig vorgenommen wurde. Diese hat eine Abteilung für Volksschulabgänger und eine höhere Abteilung für junge Leute aus höheren Schulen, welche den Berechtigungsschein erwerben wollen. Seit einigen Jahren sind für besonders tüchtige Lehrlinge der Volksschulabteilung Abendkurse eingerichtet worden, um sie in bestimmte Klassen der höheren Abteilung überzuleiten. Da der Andrang zu diesen Abendkursen zu groß wurde, erwachte das Bedürfnis, nur die wirklich über dem Durchschnitt Stehenden hinein-

[1]) Auf Grund freundlicher privater Auskünfte von Fräulein Clara Groeck, Hannover.

zunehmen. Deshalb wurden zunächst Versuche mit Prüfungen nach Moede-Piorkowski veranstaltet. Bei der Nachprüfung einer ganzen Klasse fand Beger eine sehr viel geringere Übereinstimmung der Testergebnisse mit der Schulrangordnung als Moede-Piorkowski; aber die Abweichungen selbst boten wichtige psychologische Anhaltspunkte zur Beurteilung der Schüler. In einem Falle war ein Schüler, der in der vorläufigen Klassenrangordnung den neunten Platz hatte, der weitaus beste im Test; nach einem Vierteljahr war er auch in der Klasse Primus geworden. Der Test hatte also richtig prophezeit. Bei der Auslese für die Abendkurse und bei der Überleitung in die höhere Abteilung wird jetzt „kombiniert" verfahren: es werden Zeugnisse, Beobachtungen der Lehrer, Ausfälle der pädagogischen Prüfung und Testergebnisse zur gemeinsamen Grundlage der Entscheidung gemacht. „Wenn wir wirklich im Laufe der Jahre dazu kommen müssen, eine strenge Auswahl der Begabten für die Abendkurse zu treffen, dann kann es meines Erachtens nur auf der Grundlage des Hamburger Systems sein. Im Vordergrund muß die Beobachtung durch den Lehrer stehen, die dann durch eine mäßige Anzahl von Tests ergänzt werden kann."

10. Auslese 14- und 15jähriger Mädchen zur Aufnahme in ein Lehrerinnenseminar (Hamburg).

Die folgende Auslese, über welche Penkert und Melchior berichten (*205*, *216*), vertritt einen anderen Typ als die bisher geschilderten; denn in ihr ist die psychologische Prüfung in die engste Verbindung gebracht mit der eigentlich pädagogischen Aufnahmeprüfung, ja sie ist geradezu ein Bestandteil von dieser geworden. Und vielleicht hat diese Verfahrungsweise die meiste Aussicht, in Zukunft die Auslese zu bestimmen. Freilich ist hierbei eine ziemlich weitgehende Anpassung der Methode an die Bedürfnisse des praktischen Schulbetriebes notwendig, ohne daß doch das eigentümlich Wertvolle der psychologischen Fähigkeitsfeststellung verloren gehen darf. Diese Forderung ist in den Prüfungen für das Lehrerinnenseminar noch nicht voll erfüllt worden.

Zum Eintritt in das Hamburger Lehrerinnenseminar melden sich alljährlich sehr viel mehr Bewerberinnen, als aufgenommen werden können. So betrug die Zahl der Meldungen in den Jahren 1917, 1918, 1919: 198, 130, 107. Zur Aufnahme gelangten: 27, 26, 22, also 11—20½%. Die meisten Bewerberinnen standen im Alter

von 14—15 Jahren, hatten die Volksschule oder ein Lyzeum absolviert und meldeten sich für die unterste (V.) Seminarklasse; daneben wurden auch einige Bewerberinnen für die höheren Klassen geprüft.

Früher bestand die Aufnahmeprüfung aus einem schriftlichen Examen mit Rechenaufgaben und mehreren Aufsätzen sowie einer mündlichen Prüfung. Der Zwang, die Auslese so scharf zu gestalten, führte zu dem Wunsch, auch die allgemeine geistige Fähigkeit mitzuprüfen. Die Seminarlehrer Melchior und Penkert (1919 auch Dr. P. Petersen) arbeiteten daher, beraten vom Hamburger Psychologischen Laboratorium, Testmethoden aus, die sich der Aufnahmeprüfung einreihen ließen; die anderen, rein pädagogischen Prüfmittel — Aufsätze, mathematische Arbeiten sowie mündliche Prüfungen — blieben daneben unverändert bestehen. Derartige kombinierte Aufnahmeprüfungen haben bisher viermal (1917—1920) stattgefunden. Folgende (natürlich in jedem Jahr wechselnde) Tests wurden angewandt:

1. Die Bindewortergänzung. (Ausfüllung eines Lückentextes mit fehlenden Konjunktionen nach Minkus.)

2. Der Bilderbogentest. (Die Bilder eines Münchner Bilderbogens, die eine zusammenhängende Geschichte darstellen, sind vom Text befreit worden; die Geschichte ist nebst passender Überschrift als Aufsatz zu schreiben.)

3. Definitionen. (Z. B. Vetter, Fenster, Wehrpflicht, Geld.)

4. Vergleichung von Begriffen. (Z. B. Geiz und Sparsamkeit, Arbeit und Spiel, Unterricht und Erziehung.)

5. Begriffsordnungstests. (Eine Reihe durcheinander gemischter Begriffe sollten in eine logische Ordnung gebracht werden; in einem Fall sollten die richtig geordneten Begriffe zur Grundlage eines Aufsatzes gemacht werden.)

6. Fabeltests. (Zu gegebenen Fabeln die Lehre finden.)

7. Ethische Urteilsfähigkeit. (Zu kleinen Erzählungen sollte ethisch durch Billigung oder Mißbilligung Stellung genommen werden. Nach Roth.)

8. Kritiktest. (Text mit eingestreuten Absurditäten sowie anderen — orthographischen, grammatischen — Fehlern wurde vorgelegt. Die Prüflinge hatten die Fehler zu finden und zu kritisieren.)

8. Bildbeschreibungen mit verschiedener Instruktion. (Zu einem Bild mußten die Prüflinge in verschiedener Weise Stellung nehmen, z. B. a) Darstellung der Vorgeschichte und Nachgeschichte, b) Beschreibung der Gefühle der dargestellten Personen [Einfühlungstest].)

10. Schematest. (Eine Reihe von Begriffen, die unter einen gemeinsamen Oberbegriff fallen — z. B. alle Fahrzeuge — in ein logisch geordnetes Schema zu bringen.)

Man erkennt den starken Einfluß pädagogischer Rücksichtnahmen bei dieser Auswahl z. B. daran, daß die dem Schulbetrieb

vertraute Form des Aufsatzes bei den Tests eine große Rolle spielt; noch stärker tritt jener Gesichtspunkt bei der Auswertung der Ergebnisse hervor. Die Knappheit der Zeit verhinderte eine gründliche psychologische Messung und Statistik der Resultate; und die gesamte Form der Veranstaltung erforderte, daß die Bewertung der Tests sich ganz der Bewertung der übrigen Prüfungsarbeiten einfügte. Deshalb mußten sich die Prüfer damit begnügen, für jeden Test die fünf bekannten Prädikatstufen der Schule, 1 = „sehr gut" usw. bis 5 = „ungenügend", anzuwenden; soweit irgend möglich legten sie bestimmte psychologische Beurteilungsmaßstäbe der Prädikatgebung zugrunde. (Hierbei wurden sie zum Teil durch Arbeitsgemeinschaften des Psychologischen Laboratoriums unterstützt.)

Die im Test gegebenen Prädikate wurden für jede Bewerberin mit den anderen Prüfungsprädikaten verbunden; auf Grund der schriftlichen Gesamtleistung wurde die Zulassung zur mündlichen Prüfung, auf Grund der schriftlichen und mündlichen Gesamtleistung wurde die Aufnahme ausgesprochen. Die Tests selbst hatten also im allgemeinen keinen gesonderten Anteil an der Einschätzung; doch wurde in Zweifelsfällen ihr Ausfall gern beachtet, und in Fällen ungewöhnlicher Unstimmigkeit zwischen Test- und sonstiger Leistung sah man sich zu näherer Analyse der Bewerberin angeregt.

Im großen und ganzen bestand eine erhebliche Korrelation zwischen dem Testausfall und dem Gesamtprüfungsausfall, wie die folgende Tabelle lehrt:

Aufgenommen wurden	Von allen Bewerberinnen	Von den Bewerberinnen mit guten Testleistungen	mittleren Testleistungen	schlechten Testleistungen
1917	14 %	40 %	20 %	1 %
1918	20 %	87½ %	25 %	2 %
1919	20½ %	66⅔ %	16 %	0 %

Für die aufgenommenen Prüflinge des Jahrgangs 1917 läßt sich auch jetzt schon etwas über die Erprobung sagen. Nach Penkert sind unter den 23 damals in die V. Klasse aufgenommenen Mädchen 4 bedeutend bessere Schülerinnen, als nach dem Testausfall zu erwarten gewesen wäre; unter diesen ist eine, die trotz nichtgenügender Testleistungen aufgenommen wurde und die sich nun als normal begabtes, wenn auch geistig etwas ungewandtes Mädchen erweist, die durch unermüdlichen Fleiß ihre Aufgabe be-

wältigt. Zwei andere waren auf Grund des Testausfalls überschätzt worden; es waren Mädchen von guter sprachlicher Begabung, aber ohne eigentlich tiefere Interessiertheit und Zähigkeit in der Arbeit.

Die Lehrerschaft des Lehrerinnenseminars ist der Überzeugung, daß die Einbeziehung von Fähigkeitsproben zweifellos eine Bereicherung der Prüfung und eine Sicherung der Entscheidung bedeutet. Methodisch ist hier freilich noch manches zu vervollkommnen. Vor allem muß darauf das Augenmerk gerichtet werden, ob man die spezifischen Züge der pädagogischen Begabung prüfen könne, da diese doch für jede künftige Lehrerin in erster Reihe von Wichtigkeit ist. (Vielleicht ist der Einfühlungstest — in unserer Aufzählung 9b — nach dieser Richtung auszubauen, ebenso die Idee Penkerts: über einen Bilderbogen eine Niederschrift als „Erzählung für eine 5jährige" zu fordern [*205*, S. 22]).

Näheres über die Bedeutung und die psychologischen Ergebnisse der einzelnen Tests ist in den beiden Berichten über die Prüfung für das Lehrerinnenseminar zu finden.

11. Ausleseprüfungen von Fachschülern und Lehrlingen der metallbearbeitenden Berufe.

Eine letzte Gruppe von Ausleseprüfungen kann hier nur in Kürze gestreift werden. Denn sie leitet von unserem Thema der Schülerauslese in das andere, große Gebiet der Berufseignungsprüfung über, das wir hier ausschließen müssen; auch handelt es sich nur noch zum geringeren Teil um Feststellung der Intelligenz; der Hauptnachdruck liegt auf der Untersuchung der technischen Begabung und ihrer Teilfunktionen (der konstruktiven Phantasie, des Augenmaßes, der Feinheit des Tastsinns usw.).

Die erste derartige Prüfung fand Ostern 1917 an der städtischen Gewerbeschule in Leipzig statt. Hier sollten 250 Schüler der metallbearbeitenden Berufe auf 7 Klassen verteilt werden, die den Fähigkeiten entsprechend abgestuft sind; und zwar fand die Zuteilung der Schüler auf die Qualitätsklassen statt ohne Berücksichtigung der Schulzensuren, lediglich auf Grund einiger Versuche, die der Ingenieur Professor Freund gemeinsam mit dem Gewerbelehrer Herzog an den einzelnen Schülern unternahm (*189*, *190*).

Es wurde, dem gewählten Beruf entsprechend, vor allem das Raumvorstellungsvermögen und die Fähigkeit, Ursache und Wirkung klar zu erkennen, geprüft, daneben auch die Intelligenz, und zwar mit folgenden Aufgaben:

1. Textlückenergänzung nach Ebbinghaus. 2. Entfaltungstest (aus einem mehrfach zusammengefalteten Papier wird eine Ecke herausgeschnitten; der Prüfling muß aufzeichnen, wie nach der Entfaltung das Papierblatt aussieht). 3. Umlegung von Dreiecken gegeneinander nach dem Gedächtnis. 4. Verständnis für einen kleinen einfachen Mechanismus, der keine technischen Spezialkenntnisse, sondern nur Einsicht in elementare technische Zusammenhänge voraussetzt.

Die Verschiedenartigkeit der Testleistungen war überraschend groß; die Bewertung der Antworten wurde nach möglichst exakten Maßstäben vorgenommen, zugleich aber der individuelle Eindruck des Prüfers (ob eine Antwort zögernd gegeben, ob eine von Verständnis zeugende Zwischenfrage getan wurde usw.) mit in Betracht gezogen.

Der Erfolg der Gliederung in sieben Qualitätsklassen nach diesem Verfahren war verhältnismäßig günstig; die weitere Erfahrung mit den Schülern zeigte, daß nur wenige Irrtümer vorgekommen waren. Ostern 1918 wurde ähnlich verfahren, nur daß diesmal zu den eigentlichen Tests auf Wunsch des Lehrerkollegiums Prüfungen des sprachlichen und des Rechenvermögens hinzukamen. Ergebnis: „Die Gliederung hat sich auch dieses Jahr als erfolgreich herausgestellt. Wesentliche Mängel bzw. Ungleichheiten innerhalb der Klassen sind nicht hervorgetreten. Mit den besten Klassen konnte das vorgeschriebene Pensum gründlicher und wissenschaftlicher durchgearbeitet werden als mit den minderen Klassen."

Besonders weit ausgebildet ist die psychologische Auslesemethodik bereits bei der Neuaufnahme in die Lehrlingsschulen einiger technischer Großbetriebe. Bekanntlich haben solche Betriebe, wie zum Beispiel die A. E. G., Ludwig Loewe, die Vulkanwerft in Hamburg, bedeutende Schulveranstaltungen für den technischen Nachwuchs[1]); und da der Andrang zu den technischen Lehrlingsschulen um ein Vielfaches die Zahl der Aufzunehmenden übertrifft, stellte sich das Bedürfnis eines strengen Aufnahmeverfahrens ein.

In der Werkzeugmaschinenfabrik von Ludwig Loewe in Berlin sind diese Methoden zur Auswahl des Nachwuchses für hochwertige Facharbeiter (Maschinenbauer, Dreher, Werkzeugschlosser usw.) gemeinsam von dem Leiter der Werkstättenschule Ingenieur Stolzenberg, und dem Psychologen Otto Lipmann ausgearbeitet und angewandt worden (*201, 238*).

[1]) Nach Stolzenberg gibt es in Deutschland 70 solcher Werkschulen.

Geprüft wurden die folgenden Tätigkeiten, und zwar jede in einer Reihe von Tests:

Schätzung von Raumgrößen (z. B. des Durchmessers eines Kreises, eines Bolzens).

Augenmaß (z. B. Strecken freihändig in 3 gleiche Teile teilen).

Feinheit des Tastsinns (z. B. Ordnen von Blechen nach ihrer Dicke, von Schrauben nach ihrer Gängigkeit).

Optisches Erkennen kleiner Fehler (z. B. Unebenheiten oder Abweichungen von der Rechtwinkligkeit bei prismatischen Klötzen).

Konstruktive Raumphantasie (z. B.: wie muß die Figur durch einen Schnitt zerlegt werden, damit die Teile zum Quadrat zusammengesetzt werden können? Ähnliches mit schwereren Aufgaben).

Optische Vorstellungsfähigkeit (z. B. wieviel Diagonalen hat ein Würfel? ein Fünfeck?)

Erkennen und Unterscheiden von Einzelheiten (z. B. 10 einander ähnliche Blechstücke mit je 3 charakteristisch gruppierten Löchern; welches Stück entspricht einer vorgelegten Zeichnung?)

Verständnis für ein Getriebe mit mehrfacher Riemenübertragung (z. B.: wie muß sich Rad b drehen, wenn Rad a im Uhrzeigersinne gedreht wird?)

Sorgfalt des Arbeitens (z. B. ein Etikett genau auf die Mitte eines Deckels kleben).

Natürliche Intelligenz (Schloß mit 12 Schlüsseln; möglichst schnell den richtigen finden. Der Intelligente scheidet von vornherein die unmöglichen Schlüssel aus).

Reaktionsschnelligkeit.

Die Prüfungen, die für jeden Prüfling 1½ Stunden dauern, sind bisher an etwa 500 Jugendlichen angestellt worden. Die Ergebnisse werden als sehr befriedigend bezeichnet; die Übereinstimmung zwischen dem Prüfungsausfall und der inzwischen anderweitig festgestellten Berufseignung sind nach Stolzenberg sehr weitgehend. Als Beispiel für die Brauchbarkeit der Methode erwähnt Lipmann, „daß einmal aus besonderen Gründen auch die Söhne zweier Fabrikangehörigen aufgenommen wurden, obwohl sie bei der psychologischen Prüfung schlecht abgeschnitten hatten; nach einem halben Jahre wurden alle seinerzeit geprüften Lehrlinge von Meistern gut beurteilt, mit Ausnahme jener beiden, die als untauglich bezeichnet wurden“ (*201*, S. 247).

Über die Methode der Wertung und Eignung jedes einzelnen Tests geben Lipmann und Stolzenberg eine ausführliche Darstellung (*201*), die auch zum Teil lehrreiche Vergleiche über die Leistungen von Volksschülern und Fortbildungsschülern enthält.

Als eine Zentralstelle für Lehrlingsprüfungen arbeitet neuerdings das psychotechnische Laboratorium in der technischen Hochschule in Charlottenburg unter Leitung von Moede, das teils selbst für zahlreiche Betriebe die Auslese übernimmt, teils in kurzfristigen Kursen Ingenieure und andere Nichtpsychologen mit der Handhabung der Auslesemethoden bekannt zu machen sucht. Die hierbei zur Anwendung kommenden Tests sind von Moede bisher kurz beschrieben worden (*210*); sie erstrecken sich auf: Sinnestüchtigkeit (Messen und Schätzen vermittels des Auges, Druckempfindlichkeit, Feingefühl der arbeitenden Hand, Tastsinn); Raum- und Zeitauffassung (Zusammenlegung und Zusammensetzung von Figuren, Aussuchen von Werkstücken, Geschwindigkeitsschätzung durch Auge und Ohr). Dagegen fehlen noch alle näheren Angaben über Eichung und Bewertungsprinzipien, so daß sich andere psychologische Fachleute von der Art der Durchführung des Verfahrens kein rechtes Bild machen können. Wie aber gar Nichtfachleute auf Grund der von M. bisher gegebenen kurzen Andeutungen und Probedemonstrationen zu einer auch nur einigermaßen zureichenden Anwendung der Auslese fähig sein sollen, ist schwer zu verstehen[1]). So glaubt Moede z. B. dasjenige Teilgebiet seiner Auslesemethodik, das uns hier vor allem angeht, nämlich die Prüfung der Intelligenz, in dem Zeitraum von 40 Minuten den psychologiefremden Ingenieuren so nahe führen zu können, daß sie nun zu eigener Anwendung fähig sein sollen. Auf solche Weise kann bei Fernerstehenden leicht ein unzutreffendes Bild von dem Ernst und der Schwierigkeit psychologischer Untersuchungen hervorgerufen werden. Die äußere Handhabung der Testexperimente können sich natürlich die technisch gewandten Ingenieure ohne Schwierigkeit aneignen; aber sie übersehen nur allzu leicht, daß dann — in der Deutung und Wertung der Testlösung — der eigentlich psychologische Teil der Auslese mit seiner ganzen Verantwortlichkeit und Schwierigkeit erst anfängt. Es liegt im eigensten Interesse der angewandten Psychologie, daß sie hier auf die Notwendigkeit gründlicher psychologischer Schulung hinweise und nicht einen bedenklichen Dilettantismus ermutige.

---

[1]) Vgl. die Kritik, die Roloff an dem von Moede abgehaltenen Ausbildungskurs geübt hat (*222*), und die daran sich anschließende Auseinandersetzung. Z. Ang. Ps., *16*, Heft 3—6. S. 386—393. 1920.

Literaturverzeichnis.

Das Verzeichnis enthält im allgemeinen nur solche Veröffentlichungen, auf die im Text Bezug genommen wird. (Nur für Teil E ist Vollständigkeit der deutschen Literatur angestrebt worden.) Es ist in 5 Abschnitte geteilt; der erste enthält die allgemeine Literatur, die anderen die Spezialliteratur zu den 4 Teilen des Buches.

Zur Ergänzung können folgende Literaturnachweise herangezogen werden:

a) Stern, William. Differentielle Psychologie (*12*). Literatur bis Sommer 1911.

b) Die früheren Auflagen dieses Buches:
   1. Aufl. Die Intelligenzprüfungen an Kindern und Jugendlichen. Literatur bis 1912.
   2. Aufl. Erweitert um: Fortschritte auf dem Gebiet der Intelligenzprüfung. Literatur bis 1915.

c) Meumann, E. Vorlesungen (*9*) Bd. 2. Literatur bis 1912.

d) Stern, W. u. Wiegmann, O. Methodensammlung (*8*, III). (Enthält die Literatur über Methodik und Technik der Testprüfung bis 1919.)

Abkürzungen.

Am J Ps = American Journal of Psychology. Her. Hall. Worcester, Florence Chandler.

An Ps = Année psychologique. Her. Piéron. Paris, Masson & Cie.

Ar Gs Ps = Archiv für die gesamte Psychologie. Her. Meumann †, Wirth. Leipzig, W. Engelmann.

Ar Kr = Archiv für Kriminal-Anthropologie u. Kriminalistik. Her. Gross. Leipzig, F. C. W. Vogel.

Ar Ps (f) = Archives de Psychologie. Her. Flournoy, Claparède. Genf, Kündig.

Ar N Pt (Schweizer) = Schweizer Archiv für Neurologie u. Psychiatrie.

Bh Z Ang Ps = Beiheft zur Zeitschrift für angewandte Psychologie und psychologische Sammelforschung. Her. Stern & Lipmann. Leipzig, J. A. Barth.

B Ki = Beiträge zur Kinderforschung und Heilerziehung. Her. Trüper. Langensalza, Beyer & Söhne.

Biometrica = A Journal for the Statistical Study of Biological Problems. Her. Pearson. Cambridge, University Press.

B Ps Au = Beiträge zur Psychologie der Aussage. Her. Stern. Leipzig, J. A. Barth.

Br J Ps = British Journal of Psychology. Her. Ward, Rivers. London Cambridge, University Press.

Bu Soc Ant Bruxelles = Bulletin de la Société d'Anthropologie de Bruxelles. Brüssel, Hayez.

Bu Soc Et Ps Enf = Bulletin de la Société libre pour l'étude psychologique de l'enfant. Paris, F. Alcan.

D Ps = Deutsche Psychologie. Her. Giese. Langensalza, Wendt & Klauwell.

D Sc = Deutsche Schule. Her. Rissmann †, Pretzel. Leipzig, J. Klinkhardt.

E Pd = Experimentelle Pädagogik. Her. Meumann †. Leipzig, Quelle & Meyer.

Fs Ps = Fortschritte der Psychologie und ihrer Anwendungen. Her. Marbe. Leipzig, Teubner.

J Ed Ps = Journal of Educational Psychology. Her. Bell. Baltimore, Warwick & York.

J Ps Asth = Journal of Psycho-Asthenics. Faribault, Association of Medical officers of American Institutions for Idiotic and Feeble Minded Persons.

J Ps Asth Mon = Journal of Psycho-Asthenics, Monograph Supplements.

Kl Ps = Klinik für psychische und nervöse Krankheiten. Her. Sommer. Halle, Marhold.

M Pt N = Monatsschrift für Psychiatrie und Neurologie. Her. Ziehen. Berlin, Karger.

Pd Ps Arb = Pädagogisch-psychologische Arbeiten. Veröffentlichungen des Institutus für experimentelle Pädagogik und Psychologie des Leipziger Lehrervereins. Her. Brahn. Leipzig.

P d Ps Fo = Pädagogisch-psychol. Forschungen. Her. Meumann †, Gaudig. Leipzig, Quelle & Meyer.

Pd Ma = Pädagogisches Magazin. Her. Mann. Langensalza, Hermann Beyer & Söhne.

Pd Mon = Pädagogische Monographien. Her. Meumann †. Leipzig, Nemnich.

Pd Se = Pedagogical Seminary. Her. Hall. Worcester, Florenz Chandler.

Prak Ps = Praktische Psychologie, Monatsschrift für die gesamte angewandte Psychologie für Berufsberatung und industrielle Psychotechnik. Her. Moede, Piorkowski. Leipzig, Hirzel.

Ps Bu = Psychological Bulletin. Her. Pierce. Princeton, Psychological Revieuw Company.

Ps R = Psychological Review. Her. Watson. Princeton, Psychological Company.

R Ph F = Revue philosophique de la France et de l'Etranger. Her. Ribot. Paris, Alcan.

R Ps (f) = La Revue Psychologique. Her. Joteyko. Brüssel, Misch & Thron.

Sm Ab Ps Pd = Sammlung von Abhandlungen zur psychologischen Pädagogik. Aus dem Archiv für die gesamte Psychologie. Her. Meumann †. Leipzig, Engelmann.

Wi B Pd Ps = Wissenschaftl. Beiträge zur Pädag. u. Psychol. Her. Deuchler & Katz. Leipzig, Quelle & Meyer.

Z Ang Ps = Zeitschrift für angewandte Psychologie und psychol. Sammelforschung. Her. Stern & Lipmann. Leipzig, J. A. Barth.

Z Jg Schw = Zeitschrift für die Erforschung und Behandlung des jugendlichen Schwachsinns auf wissenschaftlicher Grundlage. Her. Vogt. Jena, Gust. Fischer.

Z Ki = Zeitschrift für Kinderforschung mit besonderer Berücksichtigung der pädagog. Pathologie. Her. Trüper, Martinack, Ufer. Langensalza, Bayer & Söhne.

Z N Pt = Zeitschrift für die gesamte Neurologie u. Psychiatrie. Her. Alzheimer. Berlin, J. Springer u. Leipzig, J. A. Barth.

Zb N = Zentralblatt für Nervenheilkunde und Psychiatrie. Her. Gaupp. Leipzig, J. A. Barth.

Z Pd Ps = Zeitschrift für pädagogische Psychologie und experiment. Päd. Her. Stern, Scheibner, Gaudig, Fischer. Leipzig, Quelle & Meyer.

Z Ps = Zeitschrift für Psychologie. Her. Schumann. Leipzig, J. A. Barth.

Z Pst = Zeitschrift für Psychotherapie und medizinische Psychologie. Her. Moll. Stuttgart, Ferd. Enke.

A. Allgemeine Literatur.

1. Anschütz, G. Die Intelligenz. Osterwieck, Zickfeld, 1913, 423 S.
2. Binet, A. et Simon, Th. Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. *An Ps* 11, 1905, S. 191—244.
3. Binet et Simon. Le développement de l'intelligence chez les enfants. *An Ps* 14, 1908, S. 1—94.
4. Binet, A. Nouvelles recherches sur la mesure du niveau intellectuel chez les enfants d'école. *An Ps* 17, 1911, S. 145—201.
5. Binet, A. Die neuen Gedanken über das Schulkind. Deutsche Bearbeitung von G. Anschütz u. W. J. Ruttmann. Leipzig, Wunderlich, 1912, 289 S.
6. Bühler, K. Die geistige Entwicklung des Kindes. Jena, Fischer, 1918, 378 S.
7. Groos, K. Das Seelenleben des Kindes. 4. Aufl. Berlin, Reuther u. Reichard, 1913, 334 S.
8. Hamburger Arbeiten zur Begabungsforschung. *Bh Z Ang Ps* 18—20.
   I. Die Auslese befähigter Volksschüler in Hamburg. Bericht über das psychologische Verfahren. In Gemeinschaft mit O. Bobertag, L. Heitsch, H. Meins, M. Muchow, A. Penkert, H. P. Roloff, G. Schober, H. Werner, O. Wiegmann von R. Peter und William Stern. 1919, 157 S. Vergleiche Nr. 218 dieser Bibliographie.
   II. Untersuchungen über die Intelligenz von Kindern u. Jugendlichen. Von W. Minkus, W. Stern, H. P. Roloff, G. u. A. Schober, A. Penkert. 1919, 167 S.
   III. Methodensammlung zur Intelligenzprüfung von Kindern u. Jugendlichen. Von William Stern und Otto Wiegmann. 1920, 256 S.
9. Meumann, E. Vorlesungen zur Einführung in die experimentelle Pädagogik. 3 Bde. Das I.-Problem wird besonders in Bd. 2 behandelt. Leipzig, Engelmann, 1911—14, 2. Aufl.

10. Meumann, E. Intelligenz u. Wille. 3. Aufl. Leipzig, Quelle & Meyer, 1920, 371 S.
11. Meumann, E. Abriß der experimentellen Pädagogik. Leipzig, Engelmann, 1914, 462 S.
12. Stern, William. Die differentielle Psychologie in ihren methodischen Grundlagen. Leipzig, Barth, 1911, 503 S.
12a. Stern, William. Fragestellungen, Methoden und Ergebnisse der I.-P. I. Kongreß für Jugendbildung und Jugendkunde. Dresden, 1911. *Arb. d. Bundes für Schulreform* 5, 1912, S. 8—20.
13. Stern, William. Psychologie der frühen Kindheit. Leipzig, Quelle & Meyer, 1914, 372 S.
14. Stern, William. Die Jugendkunde als Kulturforderung. Mit bes. Berücksichtigung des Begabungsproblems. *ZPdPs* 17, 1916, S. 273—311. Auch separat. Leipzig, Quelle & Meyer, 1916, 83 S.
15. Stern, William. Die Psychologie u. der Personalismus. Leipzig, Barth, 1917, 54 S.
16. Stern, William. Die menschliche Persönlichkeit. 2. Aufl. Leipzig, Barth, 1919, 270 S.
17. Stern, William. Die menschliche Persönlichkeit u. ihr psychisches Leben. *ZPdPs* 21, 1920, S. 1—23.
18. Whipple, G. M. Manual of Mental and Physical Tests. 2 Bde. Baltimore, Warwick u. York, 1914.

B. Literatur zu Teil I.

19. Baerwald, R. Exp. Untersuchung über Urteilsvorsicht u. Selbsttätigkeit. *ZAngPs* 2, 1908, S. 338—381.
20. Claparède, Ed. La psychologie de l'intelligence. *Scientia* 22, 353 bis 368, Nov. 1917.
21. Claparède, Ed. Des diverses catégories de tests mentaux. Schweizer *ArNPt* 3 (1), 102—118, 1918.
22. Cohn, J. u. Dieffenbacher, F. Untersuchungen über Geschlechts-, Alters- u. Begabungsunterschiede bei Schülern. *BhZAngPs* 2, 1911, 213 S.
23. Hasserodt, O. Sterns Intelligenzbegriff. *Hamburgische Schulzeitung* 24 (35/36), 1916, S. 171—174.
24. Höper, W. Über den objektiven Wert von I.-P. Unter besonderer Berücksichtigung von BS. *BKi* Heft 158. Langensalza, Beyer & Söhne, 1919, 112 S.
25. Heymans, G. u. Wiersma, E. Beiträge zur speziellen Psychologie auf Grund einer Massenuntersuchung I. *ZPs* 42, 1906, S. 81—127, 258 bis 301.
26. Köhler, E. I.-P. an Anthropoiden I. *Abhdlg. der Königl. Preuß. Ak. d. Wissenschaft. Phys. Math. Kl.* 1917, 213 S.
27. Krall, K. Denkende Tiere. Leipzig, 1912, 532 S.
28. Lipmann, O. Psychische Geschlechtsunterschiede. Ergebnisse der differentiellen Psychologie. *BhZAngPs* 14, 1917, 2 Teile, 108 u. 172 S.
29. Lipmann, O. Über Begriff u. Erforschung der natürlichen Intelligenz. *ZAngPs* 13, 1918, S. 192—201.

30. Messmer, O. Zur Psychologie des Lesens bei Kindern u. Erwachsenen. *Ar Gs Ps* **2**, 1903, S. 190—298. Auch gesondert *Sm Ab Ps Pd* **1** (1), 1904.
31. Minkus, W. (†) u. Stern, William. a) Die Methodik eines Massenversuchs zur Erforschung der geistigen Leistungsfähigkeit an Volks- u. Fortbildungsschülern. b) Die Bindewortergänzung. Differentiell psychologische Ergebnisse. *Hamburger Arbeiten zur Begabungsfrschg.* II, *Bh Z Ang Ps* **19**, 1919, S. 1—71.
32. Möbius, P. J. Über den physiologischen Schwachsinn des Weibes. 9. Aufl. Halle, Marhold, 1908, 181 S.
33. Peters, W. Über Vererbung psychischer Fähigkeiten. *Fs Ps* **3** (4/6), 1915, S. 185—382.
34. Schröbler, E. Die Entwicklung der Auffassungskategorien beim Schulkinde. *Sm Ab Ps Pd* **4**, 1914, 112 S.
35. Spearman, C. „General I." Objectively Determined and Measured. *Am J Ps* **15**, 1904, S. 201—292.
36. Spearman, C. u. Krueger, F. Die Korrelation zwischen verschieschiedenen geistigen Leistungsfähigkeiten. *Z Ps* **44**, 1906, S. 50—114.
37. Spearman, C. and Hart, B. General Ability, its Existence and Nature. *Br J Ps* **5**, 1912, S. 51—79.
38. Stern, Erich. Der Begriff u. die Untersuchung der natürlichen Intelligenz. *M Pt N* **46** (4), 1919, S. 181—205.
39. Stern, William. Ein Beitrag zur differentiellen Psychologie des Urteilens. *Z Ps* **22**, 1899, S. 13—22.
40. Stern, William. Die Aussage als geistige Leistung u. als Verhörsprodukt. *B Ps Au* **1**, 1904, 147 S.
41. Stern, William. Zur vergleichenden Jugendkunde der Geschlechter. 3. Kongreß für Jugendbildung u. Jugendkunde in Breslau. *Arb. d. B. f. Schulref.* **8**, 1914, S. 17—38.
42. Stern, William. Über I.-Stadien u. -Typen beim Aussageversuch. Kritische Besprechung. *Z Ang Ps* **10**, 1915, S. 300—320.
43. Stern, Clara u. William. Die Kindersprache. 2. Aufl. Leipzig, Barth, 1920, 431 S.
44. Valentiner, Th. Die Phantasie im freien Aufsatz der Kinder u. Jugendlichen. *Bh Z Ang Ps* **13**, 1916, 168 S.
45. Ziehen, Th. Über das Wesen der Beanlagung und ihre methodische Erforschung. *Pd Ma* **683**, 1918, 32 S.

C. Literatur zu Teil II.

46. Abelson, A. R. The Measurment of Mental Ability of Backward Children. *Br J Ps* **4**, 1911, S. 268—314.
46a. Benary, W. Zur Frage der Methoden psychol. I.- u. Eignungsprüfungen. *Z Ang Ps* **17**, 1920.
47. Betz, W. Über Korrelation. *Bh Z Ang Ps* **3**, 1911, 88 S.
48. Binet, A. A propos de la mesure de l'intelligence. *An Ps* **11**, 1905, S. 69—82.
49. Binet, A. La mesure du développement de l'intelligence chez les jeunes enfants. *Bu Soc Et Ps Enf* 1911, Nr. 10 u. 11, S. 187—248.

50. Binet, A. Comment les instituteurs jugent-ils l'intelligence d'un écolier? *BuSocEtPsEnf* 1910, S. 172.
51. Bobertag, O. Über I.-P. nach der Methode B. S. I. Methodik u. Ergebnisse der einzelnen Tests. II. Gesamtergebnisse der Methode. *ZAngPs* 5, 1911, S. 105—203 u. *ZAngPs* 6, 1912, S. 495—538. Auch als Sonderausgabe erschienen. 2. Aufl. Leipzig, Barth, 1920.
52. Bourdon, B. Observations comparatives sur la reconaissance, la discrimation et l'association. *RPh* 20, 1895, S. 153—185.
53. Chotzen, F. Die I.-P.-Methode von B. S. bei schwachsinnigen Kindern. (Unter Mitwirkung von Dr. M. Nicolauer.) *ZAngPs* 6, 1912, S. 411 bis 494.
54. Cimbal, W. Taschenbuch zur Untersuchung nervöser u. psychischer Krankheiten. 3. Aufl. Berlin, 1918, 255 S.
55. Claparède, Ed. Tests de développement et tests d'aptitudes. *ArPs(f)* 14, 1914, S. 101—107.
56. Claparède, Ed. Profils psychologiques gradués d'après l'ordination des sujets. *ArPs(f)* 16, 1916, S. 70—81.
57. Décroly, O. Epreuve nouvelles pour l'examen mental et son application aux enfant anormaux. *BuSocAntBruxelles* 32, 1913, 25 S.
58. Descoeudres, Alice. Les tests de B. S. comme mesure du développement des enfants anormaux. *ArPs(f)* 14 (59), 1915, S. 225—254.
59. Deuchler, G. Über die Methoden der Korrelationsrechnung in der Pädagogik und Psychologie. *ZPdPs* 15, 1914, S. 114—131, 145—159, 229—242.
59a. Deuchler, G. Über Schlußversuche, besonders an Kindern u. Jugendlichen. *ZPdPs* 21, 1920, S. 23—41.
60. Ebbinghaus, H. Über eine neue Methode zur Prüfung geistiger Fähigkeiten u. ihre Anwendung bei Schulkindern. *ZPs* 13, 1897, S. 401—459.
61. Franken, A. Bilderkombinationen. Ein Beitrag zum Problem der I.-P. *ZAngPs* 12, 1917, S. 173—229.
62. Gassmann, E. u. Schmidt, E. Das Nachsprechen von Sätzen in seiner Beziehung zur Begabung. *WiBPdPs* Heft 3, 1913, 101 S.
62a. Giese, F. Der Schematest. *DPs* 2 (2), 1918/19, S. 92—94.
63. Goddard, H. Revision of Binet-Scala. *Training School* 8, 1911.
64. Gregor, A. I.-Untersuchungen mit der Definitionsmethode. *MPtN* 36 (1), 1914, S. 1—18.
65. Gregor, A. Untersuchungen über die Entwicklung einfacher logischer Leistungen. *ZAngPs* 10, 1915, S. 339—451.
65a. Gregor, A. Zur Bestimmung des I.-Alters mittels der Definitionsmethode. *ZKi* 25 (3/4), 1920, S. 117—137.
65b. Haberman, J. V. The Intelligence Examination and Evaluation. *Journal of the American Medical Association*, Bd. 65, 1915, S. 399—404. Ferner: The Intelligence and its Examination. *Medical Record*, May 18., 1918.
66. Habrich, J. Über die Entwicklung der Abstraktionsfähigkeit von Schülerinnen. *ZAngPs* 9, 1915, S. 189—244.
67. Heilbronner, K. Zur klinisch psychol. Untersuchungstechnik. *MPtN* 17, 1905, S. 117—132.

68. Herderschée, D. Teste für taubstumme Kinder. *Z Ang Ps* **16**, 1920, S. 40—61.
69. Hylla, E. Meumanns Vorschläge zur Ausgestaltung der Binetschen I.-P.-Methode. *Arch. f. Päd.* (Teil II, *D. päd. Forschg.*) **3**, 1915, S. 16—22, 71—80.
70. Jacobsohn. Eine Methode, um Aufschluß über das sittliche Fühlen eines Jugendlichen zu bekommen. *Z N Pt* **46**, 1919, S. 285—347.
70a. Jacobsohn. Die Fernald'sche Methode zur Prüfung des sittlichen Fühlens. *Bhft Z Ang Ps* **24**, 1920.
71. Jaederholm, G. A. Undersökningar över Intelligensmätningarnas. Teorie och Praxis. 2 Bde. Stockholm, A. Boniner, 1914, 386 u. 268 S.
72. Jaederholm, G. A. Untersuchungen über die Methode B. S. *Z Ang Ps* **11**, 1916, S. 289—340.
73. Karstädt, O. Zur Schaffung von Paralleltests. *Z Ang Ps* **13**, 1918, S. 305—353.
74. Koch, A. Exp. Untersuchungen über die Abstraktionsfähigkeit von Volksschulkindern. *Z Ang Ps* **7**, 1913, S. 332—391.
75. Köhn, K. Experimentelle Beiträge zum Problem der I.-P. *Pd Ps Fo*, 1913, 138 S.
76. Kosog, O. Suggestion einfacher Sinneswahrnehmungen bei Schulkindern. *B Ps Au* **2**, 1905/06, S. 385—400.
77. Kuhlmann, F. A Revision of the BS-System for Measuring the I. of Children. *J Ps Asth Mon* **1** (1), 1912.
78. Kuhlmann. F. The Results of Grading 1300 Feeble-Minded Children with the BS-Tests. *J Ed Ps* **4**, 1913, S. 261—268.
78a. Lehmann, H. Vergleichende Untersuchungen über die Definitionsleistungen psychisch intakter u. leicht schwachsinniger Kinder. *Z N Pt* **47**, 1919, S. 387 ff.
79. Levy-Suhl, M. Die Prüfung der sittlichen Reife u. die Reformvorschläge zum § 56 des deutschen StGB. *Z Pst* **4**, 1912. Auch separat. Stuttgart, Enke.
80. Lindworsky, J. Das schlußfolgernde Denken. Exp. psychol. Unters. Freiburg 1916, 454 S.
81. Lipmann, O. Die Wirkung von Suggestivfragen. *Z Ang Ps* **1**, 1908, S. 504—546.
82. Lipmann, O. Entwicklung der grammatisch-logischen Funktionen. *Z Ang Ps* **12**, 1917, S. 347—371.
82a. Lobsien, M. Ein Test zur Prüfung der Kritikfähigkeit. *Z Pd Ps* **20**, 1919, S. 231—234.
82b. Lobsien, M. Höhere I.-P. an Jugendlichen mit Hilfe des Bindewort-tests. *Z Pd Ps* **20**, 1919, S. 157—163.
83. Masselon, R. Ps. des déments précoces. Paris 1902.
84. Mayer, A. Über Einzel- u. Gesamtleistung des Schulkindes. *Sm Ab Ps Pd* **1** (4), 1903, 136 S.
85. Meumann, E. I.-P. bei Kindern der Volksschule in Zürich. *E Pd* **1**, 1905, S. 35—100.
86. Meumann, E. Über eine neue Methode der I.-P. u. über den Wert der Kombinationsmethoden. *Z Pd Ps* **13**, 1912, S. 145—163.

87. Meyer. Die Bedeutung der modernen Psychologie f. d. Militärwesen. *Neue Militär. Blätter* 1911, 53. Jahrg., No. 6, 9, 10.
88. Meyer. Psychologie u. militärische Ausbildung. *ZPdPs* 13, S. 81—85.
89. Netschajeff, A. Untersuchungen über die Abstraktionsfähigkeit von Schülern. *ZAngPs* 4, 1911, S. 335—346.
90. Piorkowski, C. Untersuchungen über die Kombinationsfähigkeit bei Schulkindern. *PdPsArb* 4, 1913, S. 55—91.
91. Rabinovitsch, Sophie. Resultate der experimentellen Untersuchung von Kindern nach der kurzen Methode von Rossolimo. *ZAngPs* 13, 1918, S. 210—219.
92. Ranschburg, P. Das kranke Gedächtnis. Leipzig, Barth, 1911, 138 S.
93. Riebesell, P. Untersuchungen über das Moralitätsalter. *ZPdPs* 18, 1917, S. 376—385.
94. Rieger, C. Beschreibung der I.-Störungen infolge einer Hirnverletzung nebst e. Entwurf z. e. allg. anwendbaren Methode der I.-P. *Verhdlg. d. physik.-med. Ges. zu Würzburg. N. F.* 23, Nr. 4, 1889.
95. Ries, G. Beiträge zur Methodik der I.-P. *ZPs* 56, 1910, S. 321—343.
96. Rodenwaldt, E. Zur Methode der I.-P. *ArKr* 1904, S. 235—251.
97. Rodenwaldt, E. Aufnahme des geistigen Inventars Gesunder als Maßstab für Defektprüfungen bei Kranken. *Monatsschrift f. Psychiatrie*, 17. Ergänzungsheft 1905, S. 17—84.
98. Roemer, F. Assoziationsversuche an geistig zurückgebliebenen Kindern. *FsPs* 3 (2), 1914, S. 43—101.
99. Rossolimo, G. Die psychologischen Profile. Zur Methodik der quantitativen Untersuchung der psychischen Vorgänge in normalen u. pathologischen Fällen. *KlPs* 6, 1911, S. 249—326 u. 7, 1912, S. 22—26.
100. Rossolimo, G. Berichtigungen u. Ergänzungen zur Methodik der Untersuchung der „Psychologischen Profile". *KlPs* 8, 1913, S. 185 bis 189.
101. Rossolimo, G. Zur I.-P. der Zurückgebliebenen. Eine kurze Methode. *ZAngPs* 13, 1918, S. 202—209.
102. Roth, H. Das sittliche Urteil der Jugend. Borna, Leipzig, Kommissionsverl. v. R. Neske, 1915, 78 S.
103. Saffiotti, U. L'échelle métrique de l'intelligence modifié selon la méthode Treves-Saffiotti. *AnPs* 18, 1912, S. 327—340.
104. Saffiotti, U. La misura dell' Intelligenza nei Fanciulli. Rom 1916. 286 S.
105. Sander, Hjalmar. Die experimentelle Gesinnungsprüfung. Ihre Aufgaben u. Methodik. *ZAngPs* 17, 1920, S. 59—109.
106. Schaefer, M. Elemente zur moralisch-psycholog. Beurteilung Jugendlicher *ZPdPs* 14, 1913, S. 47—59, 90—98.
107. Schultze, O. Eine neue Weise der Auswertung der Intelligenzteste. *ZAngPs* 11, 1916, S. 19—28.
108. Schober, G. u. A. Über Bilderkennungs- und Unterscheidungsfähigkeit bei kleinen Kindern. *BhZAngPs* 19, 1919, S. 94—137.
109. Schüssler, H. Ist die Behauptung Meumanns richtig: Kinder können im allgemeinen vor dem 14. Lebensjahr nicht logisch schließen? *ZAngPs* 11, 1916; 12, 1916 u. 14, 1919.

110. Sommer, R. Über die Methoden der I.-P. *KlPs* 7, 1912, S. 1—21.
111. Stern, Erich. Der Begriff u. die Untersuchung der natürlichen I. *MPtN* 46, 1919, S. 181—205.
112. Stern, Erich. Die „praktische" I. u. ihre Untersuchung. *ZPdPs* 21, 1920, S. 104—109.
113. Stern, William. Der Intelligenz-Quotient als Maß der kindlichen I., insbes. der unternormalen. *ZAngPs* 11, 1916, S. 1—17.
114. Stern, William. Über Alterseichung von Definitionstests. Eine methodol. Unters. auf Grund der Massenversuche von A. Gregor. *ZAngPs* 11, 1916, S. 90—96.
115. Stern, William. Erziehung zum psychol. Beobachten u. Denken. Ein Beitrag zur Hochschul- u. Seminarpäd. *ZAngPs* 13, 1918, S. 221—228.
116. Störring, G. Experimentelle Unters. über einfache Schlußprozesse. *ArGsPs* 11, 1908, S. 1—127.
117. Szymanski, J. S. Versuche über die Entwicklung der Fähigkeit zum rationellen Handeln bei Kindern. *ZPs* 78, 1917, S. 317—321.
118. Terman, L. M. Condensed Guide for the Use of the Stanford Revision and Extension of the Binet-Simon Measuring Scale of Intelligence 1916. (Nicht im Buchhandel.) Dieses Material ist enthalten in dem jetzt erschienenen Buch Termans, Stanford Revion. I. Source Material. II. Guide.
119. Terman, L. M. I. Suggestions for Revising, Extending and Supplementing the Binet Intelligence Tests. II. Psychological Principles Underlying the BS. Scale and Some Practical Considerations for its Correct Use. III. The Significance of Intelligence Tests for Mental Hygiene. *JPsAsth* 18, 1913—14, S. 20—23, 93—104, 119—127.
120. Torren, J. van der. Über das Auffassungs- u. Unterscheidungsvermögen für optische Bilder bei Kindern. *ZAngPs* 1, 1908, S. 189—232.
121. Treves, Z. u. Saffiotti, U. La „Scala Metrica dell' Intelligenza" di Binet e Simon. Studiata nelle Scuole communali elementari di Milano. Esposizione e Critica. *Laborat. civico di Ps. pura ed appl.* Milano 1911, 67 S.
122. Weigl, F. Experimentell-pädagog. Erforschg. der Begabungsdifferenzen. *Päd. Zeitfragen.* Donauwörth 1914, 85 S.
123. Whipple, G. M. Vocabulary and Word-buildingstests. *PsR* 15, 1908, S. 94 bis 105.
124. Winteler, J. Experimentelle Beiträge zu einer Begabungslehre. *EPd* 2. 1906, S. 1—48, 147—247.
125. Yerkes, M., Bridges, J. W. and Hardwick. A Point Scale for Measuring Mental Ability. Baltimore, Warwick u. York, 1915, 218 S.
126. Ziehen, Th. Die Prinzipien u. Methoden der I.-P. 4. Aufl. Berlin, Karger, 1918.

D. Literatur zu Teil III.

126a. Béla-Tabajdi-Kun. Die I.-P. der kriminellen Jugend u. deren Ergebnisse. *ZKi* 23 (4/5), 1918, S. 143—150.
127. Bloch, E. u. Lippa, Hedwig. Die I.-P.-Methode von BS. (1908) an schwachsinnigen Kindern. *ZAngPs* 7, 1913, S. 397—409.

128. Bloch, E. u. Lippa, Hedwig. Über Wiederholung der BS.schen I.-P. an schwachsinnigen Kindern nach einem Jahre. *ZAngPs* 9, 1915, S. 512—515.
129. Bobertag, O. Neuere Arbeiten zur I.-P (Sammelbericht über 24 Nummern). *ZAngPs* 8, 1914, S. 154—166.
130. Bobertag, O. Neuere Literatur über Tests u. I.-P. (Sammelbericht über 25 Nummern). *ZAngPs* 9, 1915, S. 309—318.
131. Brigham, C. An Experimental Critique of the BS. Scale. *JEdPs* 5, 1914, S. 439—449.
132. Burt, C. Experimental Tests of General I. *BrJPs* 3, 1909, S. 94—177.
133. Chotzen, F. Die Bedeutung der I.-P.-Methode von Binet u. Simon f. d. Hilfsschule. *Die Hilfsschule* 5 (2), 1912, 10 S.
134. Chotzen, F. Die I.-P.-Methode von BS. u. ihre Verwertung für die Schule. *ZJgSchw* 6, 1912, S. 401—458.
135. Cimbal, W. Über die Erziehung zur Rüstigkeit u. Tüchtigkeit in Schule u. Beruf u. die naturwissenschaftl. Grundl. der geistigen Begabungen. Hamburg u. Altona, Hammerich & Lesser, 1919, 112 S.
136. Decroly, O. u. Degand, J. La mesura de l'intelligence chez les enfants normaux. D'après les tests de MM. Binet et Simon, Nouvelle contribution critique. *ArPs(f)* 9, 1910, S. 81—108.
137. Desceoudres, Alice. Exploration de quelque tests d'intelligence chez des enfants anormaux et arriérés. *ArPs(f)* 11, 1912, S. 351—375.
138. Dosai-Révész, Margit. Experimentelle Beiträge zur Psychologie der moralisch verkommenen Kinder. *ZAngPs* 5, 1911, S. 272—330.
139. Dougherty, Mary, L. Report of the BS. Tests Given to 483 Children in the Public Schools of Kansas City. *JEdPs* 4, 1913, S. 338—352.
140. Elderton, W. P. Tables for testing the goodness of fit of theory to observation. *Biometrika* 1, 1901—1902.
141 Galton, F. Genie u. Vererbung. Übers. von O. u. A. Neurath. Philos.-soziol. Bücherei (Eisler) 19, Leipzig, Klinkhardt, 1910, 417 S.
142. Gilbey, W. H. ass. by Pearson, K. On the Signifiance of the Teachers Appreciation of General I. *Biometrika* 8 (1/2), 1911, S. 94 bis 108.
143. Goddard, H. Four hundred Feeble-minded Children Classified by the Binet Method. *PdSe* 17, 1910, S. 387—397.
144. Goddard, H. Two thausand Normal Children Measured by the Binet Measuring Scale of I. *PdSe* 18, 1911, S. 232—259.
144a. Hartnacke, W. Zur Verteilung der Schultüchtigen auf die sozialen Schichten. *ZPdPs* 18, 1917, S. 40—44.
144b. Hartnacke, W. Das Problem der Auslese der Tüchtigen. 2. Aufl. Leipzig, Quelle & Meyer, 1916, 71 S.
144c. Hartnacke, W. Die Ursachen der Nichterreichung der Klassenziele in den Bremer Volksschulen. *ZPdPs* 18, 1917, S. 442—456.
145. Gregor, A. u. Voigtländer, Else. Die Verwahrlosung.
I. Teil: Die Verwahrlosung der Knaben.
II. Teil: Die Verwahrlosung der Mädchen. Berlin, Karger, 1918, 585 S.
146. Hoffmann, A. Vergleichende I.-P. an Vorschülern u. Volksschülern. *ZAngPs* 8, 1913, S. 102—120.

147. Hoffmann, A. Vergleichende L.-P. an Breslauer Volks- u. Vorschülern. *Schles. Schulztg.* **42** (48/49), 1913, S. 1002—1004, 1024/25.
148. Hinckley, Alice C. The Binet Tests applied to Individuals over twelve Years of Age *J Ed Ps* **6**, 1915, S. 43—58.
149. Jeronutti, A. Applicazione della „Scala Metrica dell' Intelligenza" di BS. e dei Reattivi die Sante de Sanctis, per l'accertamento del grado dell' Intelligenza nei Fanciulli normali e deficienti. Esperienze, comparazioni e critiche. *Rivista Ped.* **3**, 1909.
150. Johnston, K. L. An English Version of B.-Tests for the Measurement of I. *Training College Record*, London, Nov. 1910.
151. Jones, H. Gertrude. On the Value of the Teachers Opinion of the General I. of School Children. *Biometrika* **7** (4), 1910, S. 542—548.
151a. Kannegießer, E. Die Bedeutung der Binet-Simonschen Altersstaffelung der Tests für die Hilfsschule. *Hilfsschule* **13** (2), 1920, S. 25—33.
152. Karstädt, O. Zeugen die Experimente der wissenschaftlichen Begabungslehre für oder gegen die Einheitsschule? *Päd. Ztg.* **42** (46), 1913, S. 837—840.
153. Karstädt, O. Die bisherigen Forschungen über die Begabungsverteilung nach sozialen Schichten. *DSc* **21** (8/9), 1919.
154. Kosog, O. Die Wertung der Testserien. (Mit bes. Berücksichtigung d. Prüfungsergebnisse bei einem 5jähr. Knaben.) *Z Ang Ps* **9**, 1915, S. 280 bis 298.
155. Kramer, F. Die L.-P. bei kriminellen u. psychopathischen Kindern. Vortrag 1911, enthalten in dem Bericht über den zweiten Verhandlungstag des I. Kongresses für Jugendbildung u. Jugendkunde zu Dresden. Okt. 1911. *Arb. d. Bundes für Schulreform* **5**, Leipzig, Teubner, 1911.
156. Kramer, F. L.-P. von abnormen Kindern. *M Pt N* **33** (6), 1913, S. 500—519.
157. Lobsien, M. L.-P. auf Grund von Gruppenbeobachtungen. *Aktuelle Fragen d. Pädag. d. Gegenwart* **8**, Langensalza, Beltz, 1914, 60 S.
158. Meumann, E. Die soziale Bedeutung der L.-P. *Z Pd Ps* **14**, 1913, S. 433—440.
159. Morlé, M. L'influence du milieu social sur le degré de l'intelligence des enfants. *Bu Soc Et Ps Enf* **12** (1), 1911.
160. Pearson, K. On the relationship of intelligence to size and shape of head, and to other physical and mental characters. *Biometrika* **5** (1906 bis 1907).
161. Pearson, K. u. Jaederholm, G. A. Mendelism and the Problem of Mental Defect. London, Dulau and Co., 1914, 47 S.
161a. Riebesell, P. Über die Beziehungen zwischen I. u. Moralität bei jugendlichen Verwahrlosten. *Z Pd Ps* **20**, 1919, S. 37—44.
162. Roloff, H. P. I.-Schätzung u. Schulrangordnung. *Bh Z Ang Ps* **19**, 1919, S. 72—93.
163. Sauerbrey, W. Über den Wert der L.-P.-Methode von BS. für klinische Zwecke. *Kl Ps* **10** (4), 1919, S. 205—233.
163a. Schmidt, H. Zur Psychologie der Leistungsreihen von Binet u. Simon (Bobertag). *Hilfsschule* **13** (1), 1920, S. 1—9.

164. Schmitt, M. Der Einfluß des Milieus und anderer Faktoren auf das Intelligenzalter. *Fs Ps* 5, 1919, S. 217—255.

165. Scott, Colin A. General I. or „School Brightness". *J Ed Ps* 4, 1913, S. 509—524.

166. Stern, Erich. Beiträge zur I.- u. Eignungsprüfung Minderbegabter. *Z N Pt* 47, 1919, S. 190—269.

167. Stern, William. Zum Vergleich von Vor- u. Volksschülern. *Z Ang Ps* 8 (1/2), 1913, S. 121—123.

168. Stern, William. Zum psychologischen Vergleich von Volks- u. Vorschülern. (Eine Erwiderung.) *Päd. Ztg.* 42 (51), 1913, S. 946/7.

169. Stern, William. I.-Schätzungen durch den Lehrer. Ihr Wert, ihre Methodik. *Z Pd Ps* 18, 1917, S. 198—206.

170. Stern, William. Verjüngung. *Z Pd Ps* 20, 1919, S. 1—12.

171. Terman, L. M., Lyman, G., Ordahl, G. u. L., Galbreath, N. and Talbert, W. The Stanford Revision of the Binet-Simon Scale and some Results from its Application to 1000 Non-Selected Children. *J Ed Ps* 6, 1915, S. 551—562.

172. Treves, Z. u. Saffiotti, U. Contributo allo studio dei rapporti tral' intelligenza e i fattori biologico-soziali nella scuole. *Riv. di Antropol.* 18 (1/2), 1913.

173. Waite, H. The Teachers Estimation of the General I. of the School Children. *Biometrika* 8 (1/2), 1911, S. 79—83.

174. Weigl, F. I.-P. von Hilfsschülern nach der Testmethode. *Z Ki* 18, 1913, S. 374—380, 455—462, 509—530. Auch separat: Beiträge zur Kinderforschung Nr. 111, 40. S.

175. Weintrob, J. u. R. The Influence of Environment on Mental Ability as shown by Binet-Simon Tests. *J Ed Ps* 3, 1912, S. 577—583.

176. Wiersma, E. D. I.-P. nach B. u. S. u. ein Versuch zur Auffindung neuer Tests. *Z Ang Ps* 8, 1914, S. 267—275.

177. Williams, J. H. A Study of 150 Delinquent Boys. *Research Laboratory of the Buckel Foundation. Bulletin* Nr. 1, Febr. 1915, 15 S.

178. Yerkes, M. u. Anderson, H. M. The Importance of social Status as indicated by the Results of the Point-Scala Method of Measuring mental Capacity. *J Ed Ps.* 6, 1915, S. 137—150.

179. Ziehen, Th. Die Geisteskrankheiten des Kindesalters einschließlich des Schwachsinns u. der psychopathischen Konstitutionen. Berlin, Reuther & Reichard, 1915.

E. Literatur zu Teil IV.

180. Adler, Martha. Mental Tests used as a Basis for the Classification of School Children. *J Ed Ps* 5, 1914. S. 22—28.

180a. Apel, M. Begabungsschulen. Vita. Deutsches Verlagshaus, Berlin 74 S. (Ohne Jahr.)

181. Battista, L. Zur Methodik der Aufnahmsprüfungen an Mittelschulen. *Volkserziehung. Nachrichten des Deutsch-österreichischen Unterrichtsamtes* 11, 1919. S. 57—65.

182. Beger, E. Förderung u. Auswahl Tüchtiger an der öffentlichen Handelslehranstalt zu Leipzig. *Zeitschrift für lateinlose höhere Schulen*, 1919, Heft 2.

183. Belohoubek, V. Begabungsproblem u. Schulorganisation. *Volkserziehung. Nachrichten des Deutsch-österreichischen Unterrichtsamtes* 11, 1919, S. 50 bis 56.

183a. Bobertag, O. Bericht über die Fähigkeitsprüfung der Schüler der Unterstufe am Arndt-Gymnasium in Dahlem. *ZPdPs* 21, 1920.

184. Brahn, M. Besinnliches zur Begabungsprüfung. *ZPdPs* 20, 1919, S. 328 bis 333.

185. Chaym, G. Experimentelle Begabtenauslese. *ZAngPs* 15, 1919, S. 419 bis 427.

186. Dannenberg, O. Auslese und Berufsberatung der künstlerisch Begabten. *PrakPs* 1 (5), 1920, S. 150—159.

186a. Eckhardt, K. u. Schüßler, H. Anleitung zum Beobachtungsbogen der Frankfurter Schulbehörden. *ZPdPs* 21, 1920.

187. Ehlers, W. A. Die Einheitsschule u. der Schulpsychologe. *Hamb. Schulztg.* 26, 1918, Nr. 10 u. 11.

187a. Engelmann, Susanne. Vergleich von Begabungsprüfung u. Lehrerurteil auf Grund einer Versuchsreihe. *ZPdPs* 21, 1920, S. 109—127.

188. Erziehungswissensch. Hauptstelle des Deutschen Lehrervereins. Erörterung über den Schulpsychologen. *Päd. Zeitg.* 46 (1), 1917, S. 8—9.

189. Freund, A. Qualitätsklassen auf Grund experimenteller Feststellungen. *Sächsische Gewerbeschule*, 1918, 7. Jahrg., Nr. 11 u. 12.

190. Freund, A. Die Gliederung nach Qualitätsklassen unter Anwendung exp. Methoden. *ZAngPs* 15 (1/2), 1919, S. 69—72.

191. Höper, W. Siehe Nr. 214 dieser Bibliographie.

191a. Höper, W. Fragen u. Ergebnisse der Schülerauslese in Hamburg. W. Gente, Hamburg 1920, 57 S.

192. Hylla, E. Der Schulpsychologe. *Arch. f. Pädag.* (I. Teil: Die pädag. Praxis) 4, 1916, Nr. 6, S. 241—252.

193. Hylla, E. Entwurf eines Fragebogens für berufspsychologische Beobachtung in der Schule. *ZAngPs* 12, 1917, S. 372—385.

194. Kaufel, R. Auswahl der Begabten unserer Göttinger Volksschulen. *Hannov. Schulztg.*, 1919, 55. Jahrg., Nr. 34, 35 u. 36.

195. Kosog, O. Der Schulpsychologe. *Schles. Schulztg.* 45, 1916, Nr. 38 u. 39.

196. Kosog, O. Nochmals der Schulpsychologe. Entgegnung. *Schles. Schulztg.* 45, 1916, Nr. 45.

196a. Kühn, W. Die experimentell-psychologische Fähigkeitsprüfung u. die Auslese der Begabten. 13. Heft der Sammlung: Deutsche Erziehung (Muthesius), Berlin 1919, 35 S.

197. Leipziger Lehrerverein, Inst. f. exp. Päd. u. Psychol. Anweisung für die psychologische Auswahl der jugendlichen Begabten. *PdPsArb* 9, 1919, 90 S.

198. Lipmann, O. I.-Messung zum Problem der schulischen Differenzierung. Methodologische u. experimentelle Beiträge. *ZAngPs* 13, 1918, S. 354 bis 391.

199. Lipmann, O. Die Mitwirkung des Psychologen beim „Aufstieg der Begabten". *Pädagog. Zeitg.* 46 (27), 1917, S. 409—411.

200. Lipmann, O. Das Zusammenwirken der Schule u. des Psychologen bei der Begabungs- und Eignungsauslese. *ZPdPs* 20, 1919, S. 153—157.

201. Lipmann, O. u. Stolzenberg, O. Methoden zur Auslese hochbefähigter Facharbeiter der Metallindustrie. *Schriften zur Ps. der Berufseignung u. d. Wirtschaftslebens,* Nr. 11, u. *ZAngPs* 16, 1920.

202. Mann, A. Der Schulpsychologe. *Schles. Schulztg.* 45, 1916, Nr. 34.

203. Mann, A. Psychographischer Fragebogen für die Breslauer Mittelschulsonderklassen. *Schles. Schulztg.* 51, 1917.

204. Martinak, E. Beiträge zum Problem der Begabung. *Volkserziehung. Nachrichten des Deutsch-österreichischen Unterrichtsamtes* 11, 1919, S. 43—50.

205. Melchior, O. u. Penkert, A. Über die Anwendung zweier psychologischer Methoden bei der Aufnahme in ein Lehrerinnenseminar (Methode der Bindewortergänzung u. der Bilderbogentest). Enthalten in Nr. 231 dieser Bibliographie.

206. Moede, W., Piorkowski, C. u. Wolff. Die Berliner Begabtenschulen, ihre Organisation u. die Methoden der Schülerauswahl. 3. Aufl. Langensalza, Beyer u. Söhne, 1919.

207. Moede, W. u. Piorkowski, C. Die Einwände gegen die Berliner Begabtenprüfungen sowie ihre kritische Würdigung. Langensalza, Beyer u. Söhne, 1919, 29 S.

208. Moede, W. u. Piorkowski, C. Zur „Bewährung der ersten Begabtenklasse" auf dem Köllnischen Gymnasium. *ZAngPs* 15, 1919, S. 434—437.

209. Moede, W. u. Piorkowski, C. Zwei Jahre Berliner Begabtenschulen. Leipzig, Hirzel, 1920. Baldiges Erscheinen angekündigt.

210. Moede, W. Die psychotechnische Eignungsprüfung des industriellen Lehrlings. *PrakPs* 1, 1919/20, S. 6—18 u. 65—81.

211. Muchow, M. Der Beobachtungsbogen. Enthalten in Nr. 218 dieser Bibliographie. Auch separat erschienen. Leipzig, Barth, 1919.

212. Muchow, Martha. Die psychol.-pädag. Methode der Auslese begabter Volksschüler für Sonderklassen mit fremdsprachlichem Unterricht. Hamburg 1918. *Preuß. Volksschullehrerinnenztg.,* 1919, Nr. 9/10, S. 66—68.

213. Muchow, Martha. Zur Frage der psychol. Beobachtung in der Schule. *Pädag. Reform* (Hamburg) 43, 1919, Nr. 45.

214. Muchow, M. u. Höper, W. Beobachtungsbogen u. Schülerauslese. *ZPdPs* 20, 1919, S. 301—315.

I. Muchow, M. Notwendigkeit u. Möglichkeit der Heranziehung des Lehrerurteils bei der Begabtenauslese.

II. Höper, W. Erfahrungen mit dem Hamburger Beobachtungsbogen 1919.

215. Muchow, Martha. Anleitung zur psychol. Beobachtung der Schüler. *Deutsche Einheitsschule* 1 (12), 1920, S. 353—358.

216. Penkert, A. Über die Anwendung von Tests bei Aufnahmeprüfungen in ein Hamburger Lehrerinnenseminar. *Hamb. Arb. z. Begabungsforschung. BhZAngPs* 19, 1919, S. 138—165.

217. Penkert, A. Vgl. Nr. 205 dieser Bibliographie.

218. Peter, R. u. Stern, William. Die Auslese befähigter Volksschüler in Hamburg. Bericht über das psychologische Verfahren. *Hamb. Arb. z. Begabungsforschung* Nr. I. *Bh Z Ang Ps* 18, 1919, 157 S.

Inhalt:

1. Überblick über das Gesamtverfahren. Rudolf Peter u. W. Stern.
2. Der Beobachtungsbogen. Martha Muchow.
3. Das Ordnen von Begriffsreihen. Heinz Werner.
4. Definition von Begriffen. H. P. Roloff.
5. Ergänzung von Textlücken. Otto Wiegmann.
6. Dreiwortmethode. Otto Bobertag.
7. Fabeltest. Lenore Heitsch.
8. Prüfung der Kritikfähigkeit. Hugo Meins.
9. Bildbeschreibung. Anton Penkert u. Gustav Schober.
10. Prüfung des sinnvollen Behaltens. Martha Muchow.
11. Anhang. Der neue Beobachtungsbogen.

219. Petersen, P. Der Aufstieg der Begabten. *Deutscher Ausschuß für Erziehg. u. Unterricht.* Leipzig, Teubner, 1916, 208 S.

220. Piorkowski, C. Vgl. Nr. 206—209 dieser Bibliographie.

221. Rebhuhn, H. Entwurf eines psychographischen Beobachtungsbogens für begabte Volksschüler. *Z Ang Ps* 13, 1918, S. 416—428. Auch separat: Leipzig, Barth, 1918.

222. Roloff, H. P. Ausbildungskursus in der Eignungsprüfung des industriellen Lehrlings, veranstaltet vom Laborat. für industrielle Psychotechnik in Charlottenburg, vom 13.—18. Okt. 1919. *Z Ang Ps* 16, 1920, S. 166—171. Die sich daran anschließende Polemik ebenda 16, S. 386—393.

222a. Schneider, F. Das Studium der Individualität. Eine notwendige Voraussetzung für die Durchführung der Schulreform und der Berufsberatung. Paderborn, Schoeningh, 1919.

223. Schönebeck, E. Erfahrungen in einer Berliner Begabtenklasse. *Deutsche Schule* 23 (3), 1919.

224. Schönebeck, E. Die Bewährung der Begabten. *Z Ang Ps* 15, 1919, S. 427—433.

225. Schönebeck, E. Ein Schlußwort zur Bewährung der Begabten. *Z Ang Ps* 15, 1919, S. 437—439.

226. Spranger, Ed. Begabung u. Studium. *Deutscher Ausschuß f. Erziehg. u. Unterricht.* Leipzig, Teubner, 1917, 99 S.

227. Steiskal, Th. Schülerbeschreibung. *Volkserziehung. Nachrichten des Deutsch-österreichischen Unterrichtsamts* 11, 1919, S. 66—74.

228. Stern, Erich. Bemerkungen zur Frage der Begabtenauslese. *Z Pd Ps* 19, 1918, S. 332—335.

229. Stern, William. Die Jugendkunde als Kulturforderung u. der „Schulpsychologe". *Schles. Schulztg.* 45, 1916, Nr. 43.

230. Stern, William. Über psychologische Begabtenprüfungen. *Pädag. Reform* 41, Nr. 31, 1917, S. 187—188.

231. Stern, William. Das psychologisch-pädagogische Verfahren der Begabtenauslese. Versuche und Anregungen. Leipzig, Quelle u. Meyer, 1918. Sonderausgabe aus *Z Pd Ps* 19 (3—5), 1918, 80 S.

232. Stern, William. Psychologische Begabungsforschung u. Begabungsdiagnose. Enthalten in Nr. 219 dieser Bibliographie, S. 105—120.
233. Stern, William. Die Methode der Auslese befähigter Volksschüler in Hamburg. *ZPdPs* 19, 1918, S. 132—143. Auch enthalten in Nr. 231 dieser Bibliographie.
234. Stern, William. Über Begabtenauslese. *Jahrb. d. Gesellschaft der Lehrmittelzentrale.* Wien, Haase, 1918.
235. Stern, William. Förderung und Auslese jugendlicher Begabungen. Vortrag. *ZPst* 7 (5), 1919, S. 291—298.
236. Stern, William. Vgl. Nr. 218 dieser Bibliographie.
237. Sterzinger, O. Aufnahmeprüfungen in die Mittelschulen im deutsch-österreichischen Unterrichtswesen. *ZAngPs* 15, 1919, S. 440—443.
238. Stolzenberg, O. Berufseignung. Vortrag, gehalten auf der Arbeitsnachweiskonferenz der Vereinigung der deutschen Arbeitgeberverbände in Lübeck. *Berichte, Heft 6 der Vereinigung deutscher Arbeitg.-Verb.* Berlin 1918.
238a. Vaerting, M. Wege zur Gewinnung neuer Maßstäbe für die Begabungs- u. Berufseignungsprüfungen. *ZPdPs* 21, 1920.
239. Weigl, F. Schulpsychologischer Dienst in größeren Städten. *Pharus. Kathol. Monatsschr.* 7, 1916, S. 927—941.
240. Weigl, F. Begründung u. Entwurf eines Erziehungsbogens. *8. Jahrbuch des Vereins für christl. Erziehungswissenschaft.*
241. Wischer, P. Zur Auswahl u. Prüfung der zeichnerisch Begabten. *ZPdPs* 20, 1919, S.
242. Wyatt, St. The Quantitative Investigation of Higher Mental Processes. *BJPs* 6, 1913, S. 101—133.
243. Wyenbergh, J. van der. Die Organisation des Volksschulwesens auf differentiell-psychol. Grundlage. Leipzig, Quelle u. Meyer, 1918.
244. Zühlsdorff, E. Das Begabungsproblem in der Grundschule. Hannover, Meyer (G. Prior), 1920, 72 S.

---

245. Schüleraufnahmen. Volkserziehung. *Nachrichten des Deutsch-österreichischen Unterrichtsamtes. Amtlicher Teil* 11, 1919, S. 93 ff.

Sachregister.

A.

Abelson-Test 80.
Abnorme Kinder 169.
Alter als Bedingung der I. 14.
Alternativtest 62.
Altersfortschritt der I. 30, 148.
Altersstaffelung 133.
Altklugheit 36.
Analogietest 120.
Anlage 12, 26.
Anpassungsfähigkeit 5.
Assoziationen 89.
Aufmerksamkeitstests 84.
Aufnahme- u. Ausleseprüfung 262.
Aussageversuch 30, 40, 92.
Auslese s. u. Schülerauslese.

B.

Bedingungen der I. 11, 242.
— äußere 15, 243.
— innere 12, 244.
Begabungs-Alters-Tests 154.
Begabungstests 151.
Beobachtung 52, 60.
Beobachtungsbogen 52, 253, 256, 258 bis 261, 278, 284, 286, 290, 292, 299.
Beobachtungsversuch 92.
Berufseignung 309.
„Bewegliche Front“ 275.
Bildbetrachtung 91.
Bilderfolgen 93.
Bilderserien (nach Heilbronner u. Schober) 95, 237.
Binet-Simon-Methode s. Staffelmethode.

D.

Definitionen 89, 104, 225, 238.
Determinierende Tendenz 31, 84.
Diagnostischer Wert der IP. 169.
Dreiwortmethode 113, 224.

E.

Eigenschaftsvergleichungen 119.
— abstrakte 120.
— konkrete 119.
Eindenkung 43, 99.
Einfühlung 43, 99.
Einheitsschule, Gliederung der 277.
Einzelprüfung 270.
Entwicklung der I. 29.
— der Methodik 46.
Entwicklungstempo 245.
Entwicklungstests 151.
Erblichkeit der I. 13, 245.
Ergänzungstests 110 (s. a. Minkustest).
Erkennungstests 94.
Erprobung 275.

F.

Fabeltest 117.
Fertigkeits- u. Fähigkeitsprüfungen 262, 290.
Fiktionsfähigkeit 95.
Finden des Wesentlichen 116.
Frühintelligenz 33.
Funktionsindex 70.

G.

Genie 5.
Gesamtintelligenz 59.
Geschlechtsunterschiede 13.
Gesetzmäßigkeit der Entwicklung 39.
Gesinnungsprüfung 98.

H.

Handfertigkeit 37.
Heilbronner-Serien siehe Bilderserien.

Hilfsschüler 173, 178.
Höhere Schulen 214, 238.

I.

Intellektualität 7.
Intelligenz, analytische 21, 27.
— Arten der 18, 27.
— Bedingungen der 11, 242.
— Bedingungen, soziale, der 226.
— Begriff der 2, 7, 205.
— Einheitlichkeit der 8, 293.
— Entwicklungsstufe der 26, 29.
— u. Gedächtnis 4.
— u. Genialität 5.
— Grade der 25.
— Innenwendung der 43.
— objektive 19, 27.
— philosophische 44.
— praktische 23, 27, 34, 37, 83.
— reaktive 18, 27, 33, 50, 272.
— u. Schulleistung 194.
— spontane 18, 27, 50, 272.
— sprachlich-logische 37.
— subjektive 19, 27.
— synthetische 21, 27.
— theoretische 23, 27, 37, 43.
— Typen der 18, 79, 99, 103, 106, 126.
— Untersuchung der 45.
— Verschiedenheiten der 18.
— Verteilung der 153, 165.
— Wachstum der 29.
— Wandlungen der 33.
— Wert der 5.
— wissenschaftliche 44.
Intelligenzalter 98, 140, 161, 196.
— eichung 57.
— forschung 53.
— funktionen, Prüfung der 70.
— maße 61, 140.
— quotient 62, 143, 162, 171, 178, 183.
— rangordnung siehe u. Rangordnung.
— rückstand 142, 169.
— schätzung 52, 63, 91, 195, 200.
— — Methodik der 205.
— — Pädagogischer Wert der 200.
— — u. Schulrangordnung 209.
— — in höheren Schulen 214.
— — — Volksschulklassen 211.
Intelligenzstillstand 144, 169, 186.
— streuung 164, vgl. auch Streuungsbreite u. Staffelstreuung.
— vorsprung 142.
— zensur 67.

K.

Kategoriale Gesetzmäßigkeit der I. = Entwicklung s. Stadiengesetz.
Kenntnisprüfungen 72.
Klassenalter 196.
Kombinationsfähigkeit 89, 110, 113.
Konvergenztheorie 17.
Konzentrationstests 84.
Korrelation 63, 65, 209, 221, 265, 293, 308.
Kritikfähigkeit 100.

L.

Lebensbewährung 279.
Lehrerurteil siehe Beobachtungsbogen Intelligenzschätzung, Vorauslese.

M.

Masselontest 113.
Massenprüfung 266.
Massenstatistik 53, 178, 293.
Mehrdimensionale Ordnungen 123.
Merkfähigkeit, Prüfung der 74.
Minkustest 31, 112, 218, 239.
Moralitätsalter 98.

O.

Ordnungstests 80, 121.

P.

Paralleltests 150.
Personalbogen s. Beobachtungsbogen.
Personalismus 258.
Profil, psychologisches, von Claparède 133.
— — — Rossolimo 130.
Prüfsysteme s. Testserien.
Psychogramm s. Beobachtungsbogen.
Pubertätszeit 43.
Punktskalamethode 67, 156.

Q.

Qualitätsziffern 65.

R.

Rangordnung 64, 195, 207, 209, 217, 266, 269, 282, 283, 289, 298, 302, 306.
Räumliches Vorstellen, Prüfung des 78.

S.

Schematest 124.
Schließen 108.
Schulbahnberatung 257.
Schulbewährung 279.
Schule als Bedingung der I. 15.
Schülerauslese 249.
— Methodisches 249.
— Bewährung der 294, 303.
— Bewertung des Tests bei der 268, 282, 283, 285, 288.
— Entscheidung bei der 271.
— Ergebnisse 291.
— Übersicht über die bisherigen 208.
Schülerpsychogramm s. u. Beobachtungsbogen.
Schulleistung 194, 198.
Schulpsychologe 58.
Schulpsychologische Ämter 252, 275.
Schulrangordnung s. Rangordnung.
Schwachsinnige Kinder 169.
Schwierigkeitsindex 69, 70.
Seltenheitswert 67, 68, 69, 90.
Sittlich Minderwertige 190.
Situationen, Verhalten zu vorgestellten 95.
Soziale Bedingungen der I. 226.
Spiel 35.
Sprachbeherrschung, Prüfung der 87.
Sprachentwicklung 37.
Sprichworttest 117.
Stadiengesetz 39.
Staffelmethode 128, 137.
— Änderungsvorschläge 144.
— Mängel der 144.
— Maßprinzipien der 140.
Staffelstreuung 141, 187.
Stellungnahme des Ich zur Welt 43.
Streuungsbreite 142, 146, 163.
Stufenordnung 121.
Suggestibilität 85.

T.

Talent 3, 204.
Technische Aufgaben, Ausführung 82.
Technisch-physikalische Verhältnisse, Beurteilung 81.
Telegrammtest 116.
Testbewertung 268, 282, 283, 285, 288.
Testrangordnung 195, 217.
Test-Schul-Vergleiche, Ergebnisse der 225.
Testserien 60, 128 (s. a. Staffelmethode).
— für Mindersinnige 151.
— psychiatrische 129.
— für die Schülerauslese 263—267, 301—302, 305, 307, 310, 311, 312.
Testzensur 269, 272.
Typen s. unter Intelligenz-Typen.

U.

Umgebungstests 153.

V.

Verdichtungsleistung 116.
Vergleichen 118.
Verschiebungstests 82.
Verstandesfragen 97.
Verteilungskurve, Gaußsche 153, 161, 163, 293.
Visualität 78.
Visualisationsaufgabe 80.
Volksschulen 211, 229.
Vorauslese 253, 255.
Vorschulen 233.
Vorstellen von Verschiebungen im Raum 81.

W.

Wandlungen der I. 33.
— qualitative 33.
Welterfassung des Jugendlichen 43.
Wesentliches, Finden des 116.
Witzverständnis 117.
Wortfindung 88.
Wortverständnis 88.

Z.

Zuordnungstests 127.
Zweiwortmethode 113.

Namenverzeichnis.

Die uneingeklammerten Ziffern geben die Seitenzahl dieses Buches an, die eingeklammerten die Nummer des Literaturverzeichnisses.

A.

Abelson, A. R. 80, (*46*).
Adler, Martha 280, (*180*).
Anderson, H. M. 230, (*178*).
Anschütz, G. 7, 10, (*1*).
Apel, M. (*180a*).

B.

Baerwald, R. 20, 27, (*19*).
Battista, L. (*181*).
Beger, E. 305, (*182*).
Belohoubek, V. (*183*).
Benary, W. (*46a*).
Betz, W. 65, (*47*).
Binet, A. 42, 46, 49, 61f., 73, 84—88, 95f., 113, 119f., 123, 128, 133—136, 141f., 147f., 157, 159, 161, 168f., 172, 179, 185, 192, 194, 196f., 201, 227, 229, 230, 264, (*2, 3, 4, 5, 48, 49, 50*).
Bloch, E. 171, 178, 185, 190, (*127, 128*).
Bobertag, O. 74f., 121, 136f., 148, 159f., 168, 170, 179, 188, 198, 223, 234, 282, (*51, 129, 130, 183a*).
Bourdon, B. 85, (*52*).
Brahn, M. 278, 305, (*184*).
Brigham, C. 159, (*131*).
Bühler, K. 33, (*6*).
Burt, C. 210, 219f., (*132*).

C.

Chaym, G. (*185*).
Childs, H. G. 117.
Chotzen, F. 155, 170f., 174, 178f., 181f., 185, 188, 192, (*53, 133, 134*).
Cimbal, W. 129, (*54, 135*).
Claparède, Ed. 11, 132, 156, (*20, 21, 55, 56*).
Cohn, J. 42, 44, 218, (*22*).

D.

Dannenberg, O. (*186*).
Decroly, O. 126f., (*57, 136*).
Degand, J. 227, (*136*).
Descoeudres, Alice 143, 171, 175, 177f., (*58, 137*).
Deuchler, G. 65, 110, (*59, 59a*).
Dieffenbacher, F. 42, 44, 218, (*22*).
Dósai-Révész, Margot 193, (*138*).
Dougherty, Mary L. 159f., (*139*).

E.

Ebbinghaus, H. 10, 59, 71, 110, 217, (*60*).
Eckhardt, K. (*186a*).
Ehlers, W. A. 275, (*187*).
Elderton, W. P. 161, (*140*).
Engelmann, Susanne 93, 103, 223, 265, (*187a*).

F.

Franken, A. 127, (*61*).
Freund, A. 309, (*189, 190*).

G.

Galton, F. 13, 158, 161, (*141*).
Gaßmann, E. 75, 77, 155, (*62*).
Giese, F. 84, 124, (*62a*).
Gilbey, W. H. 210, (*142*).
Goddard, H. 136, 159, 160f., 166, 168, 170, 185, 190f., 196f., (*63, 143, 144*).

Gregor, A. 46, 59, 105, 107, 193, (64, 65, 65a, 145).
Groos, K. 33, (7).

H.

Habermann, J. V. 136, (65b).
Habrich, J. 118, (66).
Hartnacke, W. 240, 244f., (144a, 144b, 144c).
Hasserodt, O. 10, (23).
Heilbronner, K. 95, (67).
Herderschée, D. 151, (68).
Heymans, G. 13, (25).
Hinckley, Alice C. 186, (148).
Höper, W. 7, 9, 11, 145, 261, 288, 294, (24, 191, 191a, 214).
Hoffmann, A. 233—236, (146, 147).
Hylla, E. 111, 152, 275, (69, 192, 193).

J.

Jacobsohn 99, 100, 122, (70, 70a).
Jaederholm, G. A. 136, 148f., 159, 161f., 164f., 167f., 171, 173, 178f., 229, (71, 72, 161).
Jeronutti, A. (149).
Johnston, K. L. 233, (150).
Jones H. G. 210, (151).

K.

Kannegießer, E. 173, (151a).
Karstädt, O. 119, 150, 227, 229f., 233 bis 236, 241, 243, (73, 152, 153).
Kaufel, R. 297, (194).
Koch, A. 118, (74).
Köhler, E. 33f., 83, (26).
Köhn, K. 79, (75).
Kosog, O. 86, 244, 275, (76, 154, 195, 196).
Krall, K. (27).
Kramer, F. 170f., 174, 185, 191f., 199, (155, 156).
Kuhlmann, F., 143, 170, (77, 78).
Kühn, W. (196a).

L.

Lehmann, H. (78a).
Leipziger Lehrerverein 116, 125, (197).
Levy, -Suhl, M. 98, (79).
Lindworsky, J. 109, (80).
Lipmann, O. 14, 23f., 65, 69f., 83, 86, 112, 169, 252, 261, 281, 310f., (28, 29, 81, 82, 198 199, 200, 201).
Lippa, Hedwig 171, 178, 185, 190, (127, 128).
Lobsien, M. 219, 222, (82a, 82b, 157).
Lyman, G. 162f., (171).

M.

Mann, A. 261, 275, (202, 203).
Martinak, E. (204).
Masselon, R. 113, (83).
Mayer, A. 111, (84).
Meins, H. 292, (218).
Melchior, O. 306f., (205).
Messmer, O. 20, (30).
Meumann, E. 7, 9, 22, 41, 46, 49, 76, 108f., 113f., 136, 151—155, 166, 218, 228, 233, 245, (9, 10, 11, 85, 86, 158).
Meyer 58, (87, 88).
Minkus, W. 30, 55, 112, 155f., 218, 239, (31).
Möbius, P. J. 13, (32).
Moede, W. 312, (210).
— u. Piorkowski, C. 49, 85, 91, 97, 99, 115, 120, 193, 223, 251, 265, 272, 273, 298, 300, 303f., (206, 207, 208, 209).
Morlé, M. 229, (159).
Muchow, Martha 261, 284, 288, (211, 212, 213, 214, 215, 218).

N.

Netschajeff, A. 119, (89).
Nicolauer, M. 170, (53).

O.

Ordhal, G. u. L. 162, (171).

P.

Pearson, K. 65, 158, 173, 178, 210, (160, 161).
Penkert, A. 93, 115, 288, 306f., (8, II, 205, 216, 217, 218).
Peter, R. 284, 288, 294, (218).

Peters, W. 13, 67f., (*33*).
Petersen, P. 307, (*219*).
Piorkowski, C. 66, 113f., 219, 224, (*90, 220*), vgl. auch Moede u. Piorkowski.

R.

Rabinovitsch, Sophie 131, (*91*).
Ranschburg, P. 76, (*92*).
Rebhuhn, H. 261, 299, (*221*).
Riebesell, P. 98, 193, (*93, 161a*).
Rieger, C. 129, (*94*).
Ries, G. 76, 90f., 219, 222, (*95*).
Rindfleisch 211.
Rodenwaldt, E. 73, (*96, 97*).
Roemer, F. 90, (*98*).
Roloff, H. P. 106f., 211f., 215, 225, 238, 312, (*8, II, 162, 222*).
Rossolimo, G. 79, 83, 102, 130, 131, *99*), *100, 101*).
Roth, H. 99, (*102*).

S.

Saffiotti, U. 151, 229, (*103, 104, 172*).
Sander, Hjalmar 98, (*105*).
Sauerbrey, W. 175, 184f., (*163*).
Schaefer, M. 98, (*106*).
Scheifler 211.
Schmidt, E. 75, 77, 155, (*62*).
— H. 173, (*163a*).
Schmitt, M. 229, 232, (*164*).
Schneider, E. 261, (*222a*).
Schober, G. u. A. 95, 237, (*8, II, 108, 218*).
Schönebeck, E. 298, 300, 303f., (*223, 224, 225*).
Schröbler, E. 42, (*34*).
Schüßler, H. 108f., (*109, 186a*).
Schultze, O. 67f., 151, (*107*).
Scott, C. A. 220, (*165*).
Simon vgl. Binet.
Sommer, R. 90, 129, (*110*).
Spearman, C. 8, (*35, 36, 37*).
Spranger, Ed. (*226*).
Steiskal, Th. (*227*).
Stern, Erich 25, 84, (*38, 111, 112, 166, 228*).
— Clara 37, (*43*).
Stern, William 1, 13, 17, 19f., 30, 39, 108, 143, 155, 200, 225, 233, 247, 249, 265, 275, 284, (*8. 12, 12a, 13, 14, 15, 16, 17, 39, 40, 41, 42, 43, 113, 114, 115, 167, 168, 169, 170, 218, 229, 230, 231, 232, 233, 234, 235, 236*),
Sterzinger, O. (*237*).
Störring, G. 109, (*116*).
Stolzenberg, O. 83, 310f., (*201, 238*).
Szymanski, J. S. 83, (*117*).

T.

Tabajdi-Kun, Béla 191, (*126a*).
Terman, L. M. 88, 117, 136, 139, 143, 148f., 161—168, (*118, 119, 171*).
Torren, J. van der (*120*).
Treves, Z. u. Saffiotti, U. 151, 229, (*121, 172*).

V.

Vaerting, M. 278f., (*238a*).
Valentiner, Th. 42, (*44*).
Voigtländer, Else 193, (*145*).

W.

Waite, H. 210, (*173*).
Weigl, F. 76, 91, 143, 171, 178f., 184, 190, 275, (*122, 174, 239, 240*).
Weintrob, J. u. R. 231, (*175*).
Werner, H. 288, (*218*).
Whipple, G. M. 89, (*18, 123*).
Wiegmann, O. 265, (*8, III, 218*).
Wiersma, E. D. 13, 159f., 196f., (*25, 176*).
Williams, J. H. 191, (*177*).
Winteler, J. 90, 111, 218, (*124*).
Wischer, P. (*241*).
Wyatt, St. 280, (*242*).
Wyenbergh, J. van der 275, (*243*).

Y.

Yerkes, M. 67, 151, 156f., 230f., (*125, 178*).

Z.

Ziehen, Th. 9, 10, 129, 145, 174, (*45, 126, 179*).
Zühlsdorff, E. 252, 261, (*244*).

Anhang.

Beispiel zur Berechnung einer Rangkorrelation.

(n = 23 Schüler einer Untertertia.)

Korrelation zwischen dem Klassenplatz (Nr. x) und einer Intelligenzschätzung durch den Lehrer (Nr. y).

$$\rho = 1 - \frac{6 \cdot \Sigma (\text{Nr. x} - \text{Nr. y})^2}{n\,(n^2 - 1)}$$

$$\text{wF}(\rho) = 0{,}706\,\frac{1 - \rho^2}{\sqrt{n}}$$

vgl. Differentielle Psychol. S. 303.

Schüler	Nr. x	Nr. y	Nr. x — Nr. y	(Nr. x — Nr. y)²
A	1	10	— 9	81
B	2	14	— 12	144
C	3	13	— 10	100
D	4	1	+ 3	9
E	5	5	0	0
F	6	7	— 1	1
G	7	20	— 13	169
H	8	2	+ 6	36
I	9	8	+ 1	1
K	10	15	— 5	25
L	11	9	+ 2	4
M	12	22	— 10	100
N	13	4	+ 9	81
O	14	11	+ 3	9
P	15	3	+ 12	144
Q	16	21	— 5	25
R	17	6	+ 11	121
S	18	18	0	0
T	19	12	+ 7	49
U	20	16	+ 4	16
V	21	23	— 2	4
W	22	17	+ 5	25
X	23	19	+ 4	16

$$\Sigma (\text{Nr. x} - \text{Nr. y})^2 = 1160$$

$$\rho = 1 - \frac{6 \cdot 1160}{23 \cdot (23^2 - 1)} = \mathbf{0{,}43}$$

$$\text{wF} = 0{,}706\,\frac{1 - 0{,}43^2}{\sqrt{23}} = 0{,}12.$$

VERLAG VON JOHANN AMBROSIUS BARTH IN LEIPZIG

Neue Veröffentlichungen zur Begabungsforschung und zur pädagogischen Psychologie.

Hamburger Arbeiten zur Begabungsforschung 1—3.

Zugleich Beihefte 18—20 zur Zeitschrift für angewandte Psychologie.

Das Psychologische Laboratorium in Hamburg (Direktor: Prof. W. Stern) bearbeitet neben Problemen der theoretischen Psychologie auch solche der angewandten Psychologie; es versucht, für seelenkundliche Fragen des praktischen Lebens die notwendige wissenschaftliche Vorarbeit und Mitarbeit zu leisten und steht zu diesem Zweck mit der Schulbehörde, der Berufsberatung und der Lehrerschaft in Verbindung.

Im Mittelpunkt dieser theoretischen und angewandten Arbeit steht das Begabungsproblem; die ihm gewidmeten Untersuchungen werden fortlaufend in den „Hamburger Arbeiten zur Begabungsforschung" veröffentlicht, deren erste Reihe nunmehr abgeschlossen vorliegt.

Die Studien erstrecken sich auf die Analyse der Begabungen und zwar zunächst vorwiegend auf die der Allgemeinbegabung (der „Intelligenz"), ferner auf die Differenzierung der Begabungen nach Altersstufen, Geschlechtern, sozialen und Schulbedingungen, auf die Diagnose der Begabung zu Zwecken der Schülerauslese und der Berufsberatung.

Kennzeichnend für die Methodik des Laboratoriums ist die systematische Verbindung des experimentellen und des beobachtenden Verfahrens; deshalb wird einerseits der Ausbau und die Erweiterung der Fähigkeitsprüfung durch Tests, andererseits die psychologische Vertiefung des Lehrerurteils durch Beobachtungsanleitungen gepflegt.

Nr. I. Die Auslese befähigter Volksschüler in Hamburg.

Bericht über das psychologische Verfahren. In Gemeinschaft mit O. Bobertag, L. Heitsch, H. Meins, M. Muchow, A. Penkert, H. P. Roloff, G. Schober, H. Werner und O. Wiegmann herausgegeben von **R. Peter** und **W. Stern.**

IX, 157 S. mit 2 Abb. und 6 Kurven im Text. 1919. M. 10.—

Inhalt: R. Peter und W. Stern, Überblick über das Gesamtverfahren. — M. Muchow, Der Beobachtungsbogen. — H. Werner, Das Ordnen von Begriffsreihen. — H. P. Roloff, Definition von Begriffen. — O. Wiegmann, Ergänzung von Testlücken. — O. Bobertag, Dreiwortmethode. — L. Heitsch, Fabeltest. — H. Meins, Prüfung der Kritikfähigkeit. — A. Penkert und G. Schober, Bildbeschreibung. — M. Muchow, Prüfung des sinnvollen Behaltens. — Anhang: Der neue Beobachtungsbogen.

Die Arbeit gibt eine dokumentarische Darstellung der an mehr als 1000 Volksschul-

Zu den Preisen tritt ab 1. Mai 1920 noch ein Verleger-Teuerungszuschlag von 20% sowie der Sortimenterzuschlag.

kindern vorgenommenen Auslese. Es wird hier in voller wissenschaftlicher Strenge über die gesamte Methode und ihre Ergebnisse, über Gutgelungenes und weniger Gelungenes Bericht erstattet; nur so kann für künftige Veranstaltungen ähnlicher Art ein wirklicher Nutzen gestiftet werden.

Nr. II. Untersuchungen über die Intelligenz von Kindern und Jugendlichen.

Von **W. Minkus** (†), **W. Stern**, **H. P. Roloff**, **G.** und **A. Schober**, **A. Penkert.**

IV, 167 S. 1919. M. 13.—

Inhalt: W. Minkus † und W. Stern, Die Methodik eines Massenversuchs zur Erforschung der geistigen Leistungsfähigkeit von Volks- und Fortbildungsschülern. — Die Bindewort-Ergänzung. — H. P. Roloff, Intelligenzschätzung und Schulrangordnung. — G. und A. Schober, Ueber Bilderkennungs- und Unterscheidungsfähigkeit bei kleinen Kindern. — A. Penkert, Ueber die Anwendung von Tests bei Aufnahmeprüfungen in ein Hamburger Lehrerinnenseminar. — Nachtrag zu Minkus und Stern, Methodik eines Massenversuchs.

Der Band faßt eine Reihe von Einzelstudien zusammen, die sich auf die verschiedensten Altersstufen, von der frühen Kindheit bis zur Fortbildungsschule, erstrecken. Neben dem Experiment kommt wiederum das Lehrerurteil (in dem Aufsatz über Intelligenzschätzungen) zur Geltung. Die Mitwirkung der Psychologie bei der Schülerauslese wird bei 15jährigen Bewerberinnen für das Lehrerinnenseminar geschildert.

Nr. III. Methodensammlung zur Intelligenzprüfung von Kindern und Jugendlichen.

Von **Otto Wiegmann** und **William Stern.**

IV, 256 S. mit Abb. und 3 Ausklapptafeln. 1920. M. 20.—

Inhalt: Einleitung von W. Stern. — Erster Teil: Einzeltests. — Zweiter Teil: Testzusammenstellungen. — Anhang.

Die unübersehbare Vielgestaltigkeit und z. T. schwere Zugänglichkeit der für Intelligenzprüfungen vorgeschlagenen Methoden machte ein Handbuch wünschenswert, das sie in übersichtlicher Anordnung mit genauer Gebrauchsanweisung zusammenstellt. Der Mitarbeiter am Laboratorium, Herr Lehrer Wiegmann, unterzog sich in Gemeinschaft mit dem Direktor dieser Arbeit, in der die z. Zt. erreichbare Vollständigkeit der für Kinder und Jugendliche vorgeschlagenen Intelligenztests angestrebt wird. Die Methodensammlung ist dazu bestimmt, allen auf diesem Gebiet tätigen Psychologen, Pädagogen und Aerzten als Rüstzeug zu dienen.

Als Sonderdruck aus Nr. 1 der „Hamburger Arbeiten" (Beiheft 18 zur Zeitschrift f. angew. Psychol.) erschien:

Psychologischer Beobachtungsbogen für Schulkinder. 11 S. 1919. M. —.70

Da aus Lehrerkreisen der Hamburger Beobachtungsbogen ständig angefordert wurde, so wurde er als selbständiges Heftchen dem Verkauf zugänglich gemacht; er hat sich für die Charakteristik 10jähriger begabter Kinder gut bewährt.

Zu den Preisen tritt ab 1. Mai 1920 noch ein Verleger-Teuerungszuschlag von 20% sowie der Sortimenterzuschlag.

VERLAG VON JOHANN AMBROSIUS BARTH IN LEIPZIG

Die Intelligenz der Kinder und Jugendlichen
und die Methoden ihrer Untersuchung

An Stelle einer dritten Auflage des Buches:
Die Intelligenzprüfung an Kindern und Jugendlichen

Von **William Stern**

XI, 335 Seiten mit 7 Abbildungen im Text. 1920. M. 48.—*), geb. M. 56.—*)

Inhalt: Erster Teil: Wesen, Arten und Entwicklung der Intelligenz. Zweiter Teil: Die Untersuchung der Intelligenz. Dritter Teil: Anwendungen und Ergebnisse. Intelligenz und Schule. Vierter Teil: Psychologische Methoden der Schülerauslese.

Seit längerer Zeit ist die 2. Auflage von Stern's „Intelligenzprüfungen" vergriffen. An die Stelle dieses Buches tritt nun ein neues Werk, das, seinem veränderten Titel gemäß, nicht nur die Untersuchungsmethoden, sondern auch die jugendliche Intelligenz selbst nach Wesen, Arten und Bedingungen, sowie die bisher vorliegenden Ergebnisse der exakten Intelligenzforschung behandelt. Das Buch bildet, vereint mit der „Methodensammlung" (Hamburger Arbeiten Nr. 3), eine vollständige systematische und kritische Darstellung über den heutigen Stand der Fragen in ihren theoretischen Bezügen und ihren praktischen Anwendungen.

Der 4. Abschnitt obigen Buches erscheint zugleich als Sonderausgabe unter dem Titel:

Die Psychologie und die Schülerauslese

Von **William Stern**

IV, 69 Seiten. 1920. M. 9.—*)

Da die Schülerauslese in der praktischen Neugestaltung des Schulwesens und auf der Reichsschulkonferenz eine bedeutende Rolle spielt, besteht bei Behörden und Pädagogen das Bedürfnis nach einer Schrift, die die Beteiligung der Psychologie an der Auslese grundsätzlich behandelt und zugleich erstmalig über alle Auslesen Bericht erstattet, die bisher unter psychologischer Mitwirkung vorgenommen wurden. Stern wendet sich ebenso gegen die einseitig schulmeisterliche Auffassung, die ohne Psychologie bei der Auswahl der Schüler für die höheren Schulbahnen auszukommen meint, wie gegen den einseitigen Psychologismus, der allein mit Hilfe der Experimentalprüfung über das Schicksal der Kinder verantwortlich entscheiden will; er fordert ein „pädagogisch-psychologisches" Verfahren der Auslese, bei dem psychologische Prüfung und Beobachtungsanweisung in engster Verbindung mit dem Lehrerurteil zur Entscheidung beiträgt.

Der zweite Teil aus Stern **„Die Intelligenzprüfung an Kindern und Jugendlichen"** 2. Aufl., ist auch noch einzeln zu haben. Er enthält:

Die Fortschritte auf dem Gebiet der Intelligenzprüfung 1912—1915

II, 62 S. 1916. M. 3.— + Teuerungszuschläge

*Zu den mit *) bezeichneten Preisen treten in Deutschland keine Teuerungszuschläge. Bei Lieferung ins Ausland kommt ein Valuta-Ausgleich, nach der Währung des betr. Landes berechnet, hinzu.*

Monographien über die seelische Entwicklung des Kindes.

Von **Clara** und **William Stern.** 2 Bände. Vollst. M. 36.—, geb. M. 46.—

Band I: **Die Kindersprache.** Eine psychologische und sprachtheoretische Untersuchung. 2. um ein Nachwort und eine Beobachtungsanleitung erweiterte Auflage. XII, 431 Seiten. 1920. M. 26.—, geb. M. 31.—

Band II: **Erinnerung, Aussage und Lüge in der ersten Kindheit.** 2. unveränderte Auflage. X, 160 S. m. 1 farbigen Tafel. 1920. M. 10.—, geb. M. 15.—

Pädagogisches Archiv: Die Arbeit bildet ohne Zweifel einen sehr wertvollen Beitrag zur Psychologie des Kindes sowohl durch die Mannigfaltigkeit, wie durch die sorgsame Sichtung und Verarbeitung des Materials. Sie zeigt aber auch wiederum, wie vorsichtig man sein muss in der Ableitung allgemeiner Sätze aus den Tatsachen der lebendigen Wirklichkeit und noch mehr in der Anwendung der Ergebnisse psychologischer Forschung auf die Praxis. Wo es auf das Taktgefühl des Erziehers ankommt, da kann die wissenschaftliche Psychologie nichts tun. Die Darstellung ist im allgemeinen klar und gewandt.

Die Psychologie und der Personalismus.

Von **L. William Stern.** IV, 54 S. 1917. M. 2.25

Die menschliche Persönlichkeit.

Von **William Stern.** (Person und Sache, Band II.) 2. Auflage. XVI, 270 S. 1919. M. 18.—, geb. M. 20.40

Deutsche Literaturzeitung: Den Ausführungen des Verf.s kann ich fast uneingeschränkt zustimmen. Die psycho-physisch neutrale Auffassung der Person wirft neues Licht auf eine Reihe von Grundbegriffen und Problemen unserer Psychologie. . . . Endlich wird für die Geisteswissenschaften wertvollste Vorarbeit geleistet. Mir bedeutet das Buch außerordentlich viel, eine Reformtat vor allen für unsere Psychologie. (Messer.)

Zeitschrift für angewandte Psychologie

Herausgegeben von **William Stern** und **Otto Lipmann.**

6 Hefte bilden einen Band. Der 17. Band (1920) ist im Erscheinen begriffen. Preis M. 45.—*)

Die Aufgabe der Zeitschrift ist die Bearbeitung psychologischer Probleme unter besonderer Berücksichtigung ihrer **Verwertbarkeit** für anderweitige praktische und wissenschaftliche Fragestellungen und die Ausgestaltung der besonderen experimentellen, psychographischen, statistischen und Sammel-Methoden für diese Zwecke. Hauptgebiete der Zeitschrift sind die pädagogische, Wirtschafts- und Eignungs-Psychologie, ferner die forensische, pathologische, literarische, ethnologische u. vergleichende Psychologie.

Seit 1911 erscheinen:

Beihefte zur Zeitschrift für angewandte Psychologie

Herausgegeben von **William Stern** und **Otto Lipmann.**

Die Beihefte sind einzeln käuflich. In den letzten Jahren erschienen:

Heft 12. **Jugendliches Seelenleben u. Krieg.** Materialien u. Berichte. Unter Mitwirkung der Breslauer Ortsgruppe des Bundes für Schulreform u. von O. Bobertag, K. W. Dix, C. Kik, A. Mann hrsgb. von William Stern. 181 S. m. 15 Abb. 1915. M. 7.50, geb. M. 8.70

Heft 13. Th. Valentiner. **Die Phantasie im freien Aufsatze der Kinder und Jugendlichen.** VI, 168 S. mit 1 Kurventafel. 1916. M. 8.40

Heft 14. Otto Lipmann. **Psychische Geschlechtsunterschiede.** Ergebnisse der differentiellen Psychologie. 2 Teile. IV, 108 u. 172 S. m. 9 Kurven im Text. 1917. M. 18.—

Heft 15. Franziska Baumgarten. **Die Lüge bei Kindern und Jugendlichen.** Eine Umfrage in den polnischen Schulen von Lodz. IV, 111 Seiten. 1917. M. 6.30

Heft 16. Karl Bürklen. **Das Tastlesen der Blindenpunktschrift.** Nebst kleinen Beiträgen zur Blindenpsychologie von P. Grasemann, L. Cohn, W. Steinberg. 93 Seiten mit 16 Abbildungen im Text und 6 Tafeln. 1917. M. 9.—

Heft 17. Charlotte Bühler. **Das Märchen u. die Phantasie d. Kindes.** IV, 82 S. 1918. M. 4.—

Heft 18—20. **Hamburger Arbeiten zur Begabungsforschung.**

Heft 21. **Beiträge zur Psychologie des Krieges,** Paul Plaut, Psychographie des Kriegers. Walter Ludwig, Beiträge zur Psychologie der Furcht im Kriege. E. Schiche, Zur Psychologie der Todesahnungen. 180 S. 1920. M. 13.—

Heft 22. Karl Herwagen, **Der Siebenjährige.** Versuch einer Gefühls- und Vorstellungstypik u. ihre Anwendung auf d. Gesinnungsunterricht. IV, 92 S. m. 3 Abb. 1920. M. 10.—

Heft 23. Ernst Gellhorn, **Übungsfähigkeit und Übungsfestigkeit bei geistiger Arbeit.** IV, 76 S. mit 1 Tab. 1920. M. 18.—*)

Heft 24. L. Jacobsohn-Lask. **Ueber die Fernald'sche Methode zur Prüfung des sittlichen Fühlens und über ihre weitere Ausgestaltung.** 84 S. 1920. M. 22.—*)

*Ab 1. Mai 1920 tritt zu den Preisen noch ein Verleger-Teuerungsaufschlag von 20% sowie der Sortimenterzuschlag. *) ohne Teuerungszuschläge in Deutschland*

VERLAG VON JOHANN AMBROSIUS BARTH IN LEIPZIG

Schriften zur Psychologie der Berufseignung und des Wirtschaftslebens

Herausgegeben unter Mitwirkung von

M. Brahn-Leipzig, G. Deuchler-Tübingen, J. Dück-Innsbruck, H. Herkner-Berlin, A. Wallichs-Aachen, F. Weigl-München

von

Otto Lipmann und William Stern

(Zum Teil Sonderausgaben aus: Zeitschrift für angewandte Psychologie.)

Heft 1: **Wirtschaftspsychologie und psychologische Berufsberatung.** Ein Vortrag von Dr. Otto Lipmann-Kleinglienicke. 26 Seiten. 1918. M. 1.20

„ 2: **Über eine psychologische Eignungsprüfung für Straßenbahnfahrerinnen** von Professor Dr. William Stern-Hamburg. 16 Seiten. 1918. Vergriffen!

„ 3: **Die Berufseignung der Schriftsetzer.** Bericht über eine Experimentaluntersuchung von Dr. Otto Lipmann-Kleinglienicke. **Eignungsprüfungen bei der Einführung von weiblichen Ersatzkräften in das Stuttgarter Buchdruckgewerbe,** Juli bis August 1917, von Dora Krais, Lehrerin der Frauenmeldestelle in Stuttgart. 37 Seiten. 1918. M. 1.50

„ 4: **Vorstudien über die psychologischen Arbeitsbedingungen des Maschinenschreibens** von Wilhelm Heinitz-Hamburg. 56 Seiten. 1918. M. 2.10

„ 5: **Die psychologische Analyse der höheren Berufe als Grundlage einer künftigen Berufsberatung** (nebst einem psychologischen Schema für die medizinische Wissenschaft und den ärztlichen Beruf) von Dr. med. Martha Ulrich-Berlin. 38 Seiten. 1918. Vergriffen!

„ 6: **Die Berufseignung der Kanzleiangestellten** von Prof. Johannes Dück-Innsbruck. (Nebst einer Mitteilung von Dr. Stefan von Máday.) 24 S. 1919. M. —.80

„ 7: **Psychologische Personalbogen als Hilfsmittel für Pädagogik und Berufsberatung** von Dr. N. Braunshausen-Luxemburg. 34 Seiten. 1919. M. 1.—

„ 8: **Untersuchungen über die psychische Eignung zum Flugdienst** von W. Benary, A. Kronfeld, E. Stern, O. Selz. 142 Seiten. 1919. M. 4.50

„ 9: **Die psychische Eignung der Funkentelegraphisten** von O. Lipmann. 42 Seiten 1919. M. 1.80

„ 10: **Psychologie und Verkehrswesen** von Hans A. Martens. 14 S. 1919. M. —.70

„ 11: **Methoden zur Auslese hochwertiger Facharbeiter der Metallindustrie** von O. Lipmann u. O. Stolzenberg. 77 S. 1920. M. 3.60

„ 12: **Eignungsprüfungen für Fliegerbeobachter II** von W. Benary. 62 S. mit 1 Tafel. 1920. M. 3.—

„ 13: **Fragen zur Erforschung des Wirtschaftslebens** von P. Rudolf Steinmetz. **Wirtschaftspsychologie** von Hans Henning. 52 S. 1920. M. 3.60

„ 14: **Zur Organisation der Eignungspsychologie** von Dr. Hildegard Sachs. 16 S. 1920. M. 2.80

„ 15: **Studien zur Eignungsprüfung der Straßenbahnführer.** I. Methode zur Prüfung der Aufmerksamkeit u. Reaktionsweise. Von Dr. Hildegard Sachs. (Mit einer Einleitung von W. Stern und einem Beitrag von Leidner.) 27 S. 1920. M. 2.80*)!!

Soziale Praxis: Mit den „Schriften" wird ein doppelter Zweck verfolgt: Sammlung aller einschlägigen Schriften an einer Stelle, so daß sie von allen an dem Gesamtkomplex interessierten Kreisen im Abonnement fortlaufend bezogen werden können; die einzelne Arbeit soll aber auch dem besonderen Interessentenkreise durch Einzelvertrieb zugänglich sein. — In der Zeit des Wiederaufbaues nach dem Kriege wird es mehr als je darauf ankommen, die richtige Arbeitskraft an die richtige Stelle zu bringen. Die Nutzbarmachung der psychologischen Forschung für die Berufswahl und Berufsauslese kann daher noch recht bedeutungsvoll werden.

*Außer dem Verleger-Teuerungszuschlag tritt zu vorstehenden Preisen noch der Sortimenter-Aufschlag hinzu. *) ohne Teuerungszuschläge in Deutschland.*

Dr. Otto Lipmann-Berlin-Kleinglienicke.

Grundriß der Psychologie für Pädagogen

VI, 100 Seiten. 1909. M. 3.—, geb. M. 4.20

Allgemeine Deutsche Lehrerzeitung: . . . Das Wesentlichste jeden einzelnen Problems wird scharf herausgestellt und ein Zusammenhang mit dem Ganzen der Psychologie aufgewiesen; die pädagogischen Folgerungen werden sodann knapp und klar angereiht. Ich darf dieses Buch jedem in die Hand wünschen, der sich einen Ueberblick über die neuere pädagogische Psychologie verschaffen will.

Zeitschrift für Philosophie und Pädagogik: Trotz der Knappheit leidet nirgends die Klarheit; die Darstellung ist durchaus schlicht und einfach. Das Buch löst seine Aufgabe in vorzüglicher Weise.

Grundriß der Psychologie für Juristen

Mit einem Vorwort von Geheimrat Prof. Dr. Franz v. Liszt

2. veränderte und vermehrte Auflage

VII, 95 Seiten. 1914. M. 4.50, geb. M. 5.70

Archiv f. Kriminal-Anthropologie: . . . Ich halte es für selbstverständlich, dass jeder Kriminalist diese ausgezeichneten Vorträge in dem vorliegenden Buche studiert.

Zeitschrift für Psychologie: Aus Vorträgen, deren Form auch geblieben ist, hervorgegangen, muss dieses Schriftchen als ein wertvolles Hilfsmittel zur Einführung der Juristen in die Elemente der Psychologie bezeichnet werden. Jeder juristische Praktiker sollte vor weiterer psychologischer Lektüre das Büchlein lesen.
Julius Friedrich-Giessen.

Deutsche Juristenzeitung: Hierzu sind die Arbeiten wie die vorliegende ausgezeichnete Behelfe. Ursprünglich waren es Vorträge im von Lisztschen Seminar, und so sind diese in Buchform jedem Kriminalisten, der es mit seinem Berufe ernst meint, zugänglich. Ich empfehle das Buch als Vorschule dringend.
Prof. Dr. H. Gross-Graz.

Psychische Geschlechtsunterschiede

Ergebnisse der differentiellen Psychologie.

2 Teile. IV, 108 und 172 Seiten mit 9 Kurven. 1917. M. 18.—

Frauenbildung: Der Verf. geht mit ausserordentlicher Vorsicht und Genauigkeit vor, so dass die Auseinandersetzungen über die Grundsätze seiner Statistik an sich schon höchst beachtenswert sind. Die Tabellen und Kurven, welche der statistischen Arbeit des Verfassers zugrunde gelegen haben, werden im 2. Teile mit allen Einzelheiten so genau mitgeteilt, dass sie für jede weitere Bearbeitung der allgemein psychologischen und speziell schultechnischen Frage gebraucht werden können.

Literarisches Zentralblatt: . . . Es ist daher auf das lebhafteste zu begrüßen, dass sich der bewährte Mitherausgeber der „Zeitschrift für angewandte Psychologie" der weitsichtigen mühevollen Arbeit unterzogen hat, die zerstreuten Ergebnisse von nahezu 5000 Einzeluntersuchungen durch sinnvolle, zuverlässige Umwertung auf vergleichbare Basis zu bringen.

Psychologie für Lehrer.

VII, 196 Seiten. 1920. M. 9.60, geb. M. 12.—

Lit. Zentralblatt für Deutschland: Es war ohne Frage ein glücklicher Gedanke, diejenigen Ergebnisse der Psychologie zusammenzustellen, die geeignet sind, die Methodik des Unterrichts und der Erziehung wissenschaftlich zu begründen. Daß Lipmann sich dieser Aufgabe geschickt entledigen würde, war zu erwarten. Daher kann dieses aus psychologischen Lehrgängen für Lehrer und Seminaristen hervorgegangene, der Arbeitsgemeinschaft für exakte Pädagogik im Berliner Lehrerverein und den Berliner Kriegsseminaristen gewidmete Buch angelegentlichst empfohlen werden.

Preuss. Fortbildungs-Zeitung: Der Verfasser hat als Leiter der „Arbeitsgemeinschaft für exakte Pädagogik" im Berliner Lehrerverein, der er auch das Buch gewidmet hat, das Bedürfnis der Lehrer an psychologischen Kenntnissen, sowie die Verbindung zwischen psychologischem Wissen und seiner Verwendung in der Schulpraxis kennen gelernt und damit die Grundlage für sein Werk gefunden. Wissenschaft und Praxis haben sich gegenseitig befruchtet. Er bietet in knappem Rahmen eine vollständige Psychologie, behandelt aber die Abschnitte ausführlicher, die der Begründung unterrichtlicher und erziehlicher Maßnahmen dienen oder zum Verständnis der seelischen Eigenarten nötig sind. Alle Abschnitte stehen miteinander in innerer Verbindung, so daß das Ganze den Eindruck der Geschlossenheit macht. Darin liegt ein Vorzug des Buches, das nach Vorlesungen entstanden ist und deshalb die sprachliche Form des Vortrages zeigt. Das Buch ist in hervorragender Weise geeignet, seinem Zwecke zu dienen, und kann besonders auch den angehenden Lehrern und Lehrerinnen der Berufsschulen aufs wärmste empfohlen werden. Aber auch „erfahrene" Lehrer und Lehrerinnen werden es mit Erfolg studieren.

Zu den Preisen tritt ab 1. Mai 1920 noch ein Verleger-Teuerungszuschlag von 20%, sowie der Sortimenterzuschlag.

Hundert psychologische Schulversuche mit Angabe der Apparate

zusammengestellt von

Prof. Dr. A. Höfler und **Prof. Dr. St. Witasek**

Vierte Auflage

X, 53 Seiten mit 18 Abbildungen, 1 farbigen Tafel und einer Zugabe:

Weitere 25 Versuche

zusammengestellt von **Dr. Willibald Kammel**

1918. M. 3.20

Deutsche Blätter für erziehenden Unterricht: Die „Hundert Schulversuche" können für die erste Einführung in den Psychologie-Unterricht warm empfohlen werden.

Zeitschrift für Psychologie: Die neue Auflage dieses Buches, das wohl das einzige deutsche ist, das speziell für den Anfangsunterricht in der experimentellen Psychologie geschrieben wurde und diesen Zweck so ausgezeichnet erfüllt, hat die gleiche Anzahl Versuche beibehalten, allerdings mit einigen Aenderungen in der Art derselben. Eine wesentliche Verbesserung haben die Illustrationen erfahren: es sind mehrere neue hinzugekommen und viele der alten Auflage sind durch schönere ersetzt worden.

Die Elemente der Erziehungs- und Unterrichtslehre

auf Grund der Psychologie der Gegenwart dargestellt

von **Prof. Dr. Paul Barth**

Siebente und achte, wiederum durchgesehene und erweiterte Auflage.

Die Ausgabe erfolgt im Herbst 1920. Etwa VII, 728 Seiten. 1920. Etwa M. 50.—

Die Deutsche Schule: Das Werk, das sich durch Klarheit und Einfachheit der sprachlichen Form auszeichnet, verdient nachdrückliche Empfehlung. Es ist wertvoll für den erfahrenen Pädagogen wie für den Neuling und Anfänger. Es gibt nicht bloss allgemeine Richtlinien, sondern auch Einführung in die wissenschaftliche Forschung pädagogischer Einzelfragen, beachtenswerte Ratschläge für die Praxis und durch die reichen Literaturangaben treffliche Hinweise für die Weiterbildung,

Die Mittelschule: Dieses Buch darf man wohl mit Recht als die beste systematische Pädagogik der Gegenwart bezeichnen.

Neue Jahrbücher für das klassische Altertum, Geschichte usw.: Nicht Offenbarungen eines Projektenmachers oder rein praktische Erfahrungen bietet uns das Buch, sondern abgeklärte wissenschaftliche Ergebnisse, welche sich aus kulturgeschichtlicher Forschung, aus philosophischer Denkarbeit, aus schulmännischer und experimenteller Erfahrung zusammensetzen. . . . Wer sich auf dem Gebiete der Pädagogik eine wissenschaftlich begründete Meinung verschaffen will, wird sich fernerhin an Barths Buch halten müssen. Den Lehrer erhebt ein derartiges Buch auf eine Höhe der Betrachtung, die ihn beglückt.

Kinderpsychologie und experimentelle Pädogogik

von

Prof. Dr. Ed. Claparède

Nach der 4. französischen Ausgabe

übersetzt von Franz Hoffmann

XV, 340 Seiten mit 13 Figuren. 1911. M. 7.20.

Zentralblatt für Physiologie: Dieses Buch, bietet in knapper, sehr übersichtlicher Darstellung eine vorzügliche Beschreibung der Hauptkapitel der praktischen Psychologie, soweit sie die Kenntnis der Entwicklung des Kindes vermittelt. Es gewährt dem Leser nicht nur Belehrung, sondern auch Anregung zum weiteren Nachdenken über die interessanten Probleme und fördert die weitere Forschung. Sehr vorteilhaft ist in dieser Hinsicht auch die ausführliche Literaturübersicht.

Deutsche Juristenzeitung: Die Darstellung zeigt die bei Claparède bekannte Eleganz und Klarheit. Die Uebersetzung liest sich gut.

Zu den Preisen tritt ab 1. Mai 1920 noch ein Verleger-Teuerungszuschlag von 20% sowie der Sortimenterzuschlag.

VERLAG VON JOHANN AMBROSIUS BARTH IN LEIPZIG

Über Intelligenzprüfungen

nach der Methode von Binet und Simon

von **Dr. Otto Bobertag**

Sonderabdrucke von Aufsätzen aus Band 3, 5 u. 6 der Zeitschrift für angewandte Psychologie u. psychologische Sammelforschung herausgeg. von William Stern und Otto Lipmann

2. unveränderte Auflage. 176 S. 1920. M. 18.—*)

Da die Hefte der Zeitschrift, welche die Abhandlungen enthalten, von verschiedenen Seiten einzeln verlangt wurden, hat sich ein Neudruck der Abhandlungen in zusammenhängender Form als wünschenswert erwiesen.

Entwicklung der Seele des Kindes

nebst kurzer Charakteristik der Psychologie des reiferen Alters

von

Prof. Dr. A. J. Sikorsky

Zweite, vermehrte und verbesserte Auflage

VI, 159 Seiten mit 16 Abbildungen. M. 4.80, geb. M. 6.90

Zeitschrift für Kinderpflege: Sikorsky hat in seinem ausserordentlich interessanten Buche die seelische Entwicklung des Kindes, besonders des Säuglings in einer Weise durchgeführt, die sofort den Psychologen und Psychiater und nicht in letzter Linie den Kinderfreund herausblicken lässt. . . . Vier Typen der kindlichen Individualität dieses Alters sind besonders festzustellen, der beschreibende, beobachtende, emotionelle und formale Typ. Mit einer kurzen Charakteristik der Psychologie des reiferen Alters schliesst das Buch, das Eltern, Lehrern, Psychologen und Aerzten eine tiefe Fundgrube für die Erkennung der Kindesseele sein wird.

Monatsschr. für höhere Schulen: Das Buch kann aufs wärmste allen Erziehern empfohlen werden.

Die psychologische Methodologie der wirtschaftlichen Berufseignung

Von

Dr. Curt Piorkowski

2. vermehrte und bis zum gegenwärtigen Stande fortgeführte Auflage. 1919. M. 7.20

Bildet Beiheft 11 zur Zeitschrift für angewandte Psychologie

Literarischer Jahresbericht: Den Stand der unvergleichlich wichtigen Frage, wie die Berufswahl mit Hilfe psychologischer Arbeit erleichtert werden kann, lernt man aus Piorkowskis gründlicher Arbeit kennen.

Archiv für Sozialwissenschaft und Sozialpolitik: Wenn sich jemand über die auf die Berufseignung angewandte Psychologie nicht aus den Quellen, sondern aus einem Referat orientieren will, so ist Piorkowskis Arbeit sehr geeignet. Sie ist einerseits ein kritisches Übersichtsreferat, andererseits eine Art Programmschrift. Piorkowski bemüht sich, aus der großen Zahl „exakter" psychologischer Methoden jene herauszustellen, die sich wirklich praktisch verwenden lassen, die nicht lebensfremd bleiben, sondern die im technischen Betriebe, in der Schule oder der Berufsberatungsstelle angewandt werden können.

Entwurf eines psychographischen Beobachtungsbogens für begabte Volksschüler

(Aus der Arbeitsgemeinschaft für exakte Pädagogik, Berliner Lehrerverein.)

Von

Rektor **Hermann Rebhuhn** in Berlin

(Sonderabdruck aus der Zeitschrift für angewandte Psychologie, Band 13)

15 S. 1919. M. —.45

*Zu den Preisen tritt ab 1. Mai 1920 noch ein Verleger-Teuerungszuschlag von 20%, sowie der Sortimenterzuschlag *) ohne Teuerungszuschläge in Deutschland.*

G. Pätz'sche Buchdr. Lippert & Co. G. m. b. H., Naumburg a. d. S.

FEB 9 1922

STERN

INTELLIGE

DER

KINDER

3

BF
431
.S84
1920

UNIV
OF
MICH